The automatic classification of the modulation type of communication signals

Sapiano, Philip Charles

Award date:
1997

Awarding institution:
University of Bath

The Automatic Classification of the Modulation Type of Communication Signals

Submitted by Philip Charles Sapiano
for the degree of PhD
of the University of Bath

1997.

P. C. Sapiano

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary

Modulation classification has been approached using non-deterministic and decision theoretic techniques. The majority of the published literature describes non-deterministic methods where many of the classifiers decompose the signal into the instantaneous envelope, phase and frequency. A wide variety of ad-hoc techniques can then translate these signals into features for classification.

A number of decision theoretic techniques have been developed for modulation classification, which apply statistical methods for pattern recognition. These are seen to be well suited to digitally modulated signals, which are generated in a deterministic fashion.

The thesis has focused on decision theoretic methods for PSK classification in additive Gaussian white noise, the foundations of which are based upon mathematical models. Such models are developed and may be applied directly to decision theoretic modulation classification, but may also be extended to provide a greater understanding of the feature extraction stage of non-deterministic modulation classification.

A number of these mathematical models have been developed for the decision theoretic classification of PSK signals, where statistical algorithms are used to determine the number of PSK states present in additive Gaussian white noise.

Existing decision theoretic techniques for PSK classification are examined and certain methods are further developed. The main contribution however, is the development of a number of new techniques, which have been divided into two categories. Firstly the carrier synchronous techniques which have been examined, in part within the published literature and secondly the asynchronous techniques which have not.

The thesis develops two novel synchronous phase based techniques, namely the "DFT of Phase Histogram" and "Maximum Likelihood DFT of Phase Histogram". A globally optimum technique, the "Maximum Likelihood IQ Technique" is developed based on the received I and Q samples, and this method is found to have globally optimum performance. However, the different techniques require different a-priori information, which has a significant effect on the classifier implementation.

Classification performance therefore, is not the only consideration in the choice of a modulation classifier. The decision theoretic model is developed within a paradigm of assumptions, and errors in the estimation of a-priori parameters will cause the model to

become inaccurate. A further consideration is the sensitivity of the techniques to estimation errors and paradigm degradation, and these are examined within the thesis.

The new carrier synchronous techniques which have been developed in this thesis are found to perform well compared to existing methods in terms of parametric sensitivity, classification performance, computational complexity and required a-priori information.

Three asynchronous techniques have been developed which do not require carrier synchronisation and are therefore much simpler to implement than the synchronous techniques. All three techniques are based on the phase difference between adjacent symbols, and are generally of low complexity. However, there is a substantial degradation in classification performance when the methods are compared with carrier synchronous techniques.

Contents

Glossary

ϕ ..... Phase

Ω ..... Phase Difference

ψ ..... Phase Difference Modulo 2π

σ ..... Standard Deviation

λ_m ..... Log Likelihood Ratio

A/D ..... Analogue to Digital

AGWN ..... Additive Gaussian White Noise

AM ..... Amplitude Modulation

b_n ..... PSK Fourier Coefficient

BPSK ..... Two Level Phase Shift Keying

cs ..... Synchronous Phase Reference

CW ..... Carrier Wave

DFT ..... Discrete Fourier Transform

DSB ..... Double Side Band

DSP ..... Digital Signal Processing

FFT ..... Fast Fourier Transform

FIR ..... Finite Impulse Response

FM ..... Frequency Modulation

FSK ..... Frequency Shift Keying

IEE ..... Institute of Electrical Engineers

IEEE ..... Institute of Electrical and Electronic Engineers

L ..... Number of Samples

M ..... PSK Number

MLP ..... Multi-Layer Perceptron

MSK ..... Minimum Shift Keying

N ..... Number of Histogram Bins

ns ..... Non-Synchronous Phase Reference

Pdf ..... Probability Density Function

PSK ..... Phase Shift Keying

QAM ..... Quadrature Amplitude Modulation

qLLR ..... Quasi Log Likelihood Ratio

QPSK ..... Four Level Phase Shift Keying

R.F. ..... Radio Frequency

SC ..... Suppressed Carrier

SNR ..... Signal to Noise Ratio

SSB ..... Single Sideband

z ..... Envelope

Chapter 1 : Introduction

1. Introduction

In order to transmit information across a radio channel, the information bearing signal is encoded into a modulation format, and is then transmitted at the desired radio frequency. There are applications where it is useful to determine automatically the form of the transmitted modulation scheme.

The work undertaken in this thesis looks into the automatic classification of the modulation type for communication signals. This forms the basis of a pattern recognition problem, where a signal presented to the system will be classified as one of a library of modulation formats. The complexity of the pattern recognition task increases and the reliability decreases as the number of candidate schemes is increased. There are a wide variety of modulation schemes which may be applied to an information bearing signal, and a practical classifier will generally take a limited subset of these for classification purposes in order to reduce the complexity of the system.

The task of modulation recognition suffers from drawbacks when compared with typical communication detection techniques due to the lack of a-priori information. As an example, a communications system may be transmitting four level PSK at 64 kbits/s and at around 600MHz. The receiver will have knowledge of this, and so will apply a fourth power law to the signal. It will then search for a carrier component around four times the I.F. frequency, bandpass filter the signal and noise, and coherently demodulate the signal. This will then be followed by equalisation and matched filter detection. A modulation recognition system is unable to do the same, as it has no prior knowledge that the signal is digitally phase modulated, there is no knowledge of the symbol shaping, no equalisation techniques are available and there is no knowledge of the bit rate or carrier frequency of the signal.

This lack of a-priori information makes the modulation recognition problem a complicated task, and when noise, interference, distortion and multi-path effects are added, the problem becomes even more difficult to solve.

The modulation schemes considered within this work are frequency domain separated signals, which are the most common forms of modulation. Transmissions of these types may be found across the range of radio frequencies, although some modulation formats may be more popular in a particular frequency band due to frequency dependent propagation effects.

1.1 Applications

The primary applications of modulation recognition fall into the following categories :

A) Military and Civil Surveillance

In military and civil surveillance applications, modulation recognition is used as an aid in the monitoring of enemy transmissions and the identification of a threat. In general, the modulation type of the signal is one parameter in a complex problem. The significance of the transmission will depend upon a set of parameters including the modulation type. Other parameters may include the direction and frequency of transmission.

Examples of the ultimate goal are to either decode the information on the transmission or to jam a signal which may be linked to a terrorist detonator.

Modulation recognition has been conventionally achieved by manual means, with the aid of various equipment items such as different demodulators, an oscilloscope for time domain representation, a spectrum analyser and a modulation meter. This approach requires a human operator to continually monitor the radio channel, which is costly and for fast burst transmissions can be difficult to process effectively. One human operator cannot listen in to all the demodulated outputs simultaneously, so there will be a time delay associated with the switching between demodulators.

Whilst an automatic system is capable of improving the performance in certain areas, it is not always possible for it to outperform the discriminating capabilities of a human. This is emphasised in the classification of analogue signals, where a human will quickly and effectively be able to detect the presence of speech at the output of a correct demodulator. This is an extremely difficult process for an automatic system, and is illustrated by the complexity and lack of reliability of speech recognition algorithms in noisy environments.

In the case of digital transmissions, the signals may be more easily modelled and the discriminating capabilities of a human will not be as significant in these cases.

B) Spectrum Management & Interference Identification

It is essential to have tight control of transmissions over the radio spectrum in order that user transmissions are reliable. It is therefore important to be able to detect that transmissions have been approved by the relevant licensing authorities.

In policing the airwaves, it is useful to be able to determine the modulation type of a signal being transmitted to determine if the transmission is authorised.

C) Radio Direction Finding

In radio direction finding systems, a particular signal is tracked. One descriptive property of a signal being tracked is the modulation type of the signal. Using modulation recognition, a user may determine when an incorrect transmission is being tracked.

D) Modulation Diverse Communication Systems

There are applications in mobile radio which enable the use of a range of modulation types. Two variants of this are in a variable rate digital communications system and the gradual migration from an analogue set of mobile radios to a digital platform.

Variable rate communication systems will adaptively alter the modulation scheme depending upon the channel characteristics. In a good noise free channel a large number of bits per symbol will be transmitted, and in a poor environment fewer bits will be transmitted. Modulation recognition is not a primary task in this form of system as the switching of the modulation scheme will be protocol driven. However, if a user is required to observe the information part way through the transmission, or if modulation synchronisation becomes lost, modulation recognition may provide a useful technique.

In private mobile radio infrastructures where users are upgrading from analogue to digital systems, it is often desirable to upgrade the mobile radios as a gradual process and spread the capital outlay. In such systems it would be desirable to enable the base-station and mobile units to have the capability of switching between analogue and digital transmissions according to what has been received.

The modulation recognition in this type of application tends to be considerably more simple than that of the military style of modulation recognition as the number of schemes will be considerably limited, and there will be additional a-priori information available.

E) Intelligent Commercial Receiver Equipment

Modulation recognition may be applied to intelligent receivers. One example is a commercial scanning receiver, which enables a user to eavesdrop on communications. Another example is within communication measurement equipment such as a communications analyser, which is generally switched manually to cater for the different modulation types.

1.2 Modulation Recognition Techniques

The general procedure for modulation recognition is shown in figure 1.1. The radio frequencies are scanned for a signal of interest and once a signal is detected, it is translated to an I.F. where a segment is captured digitally.

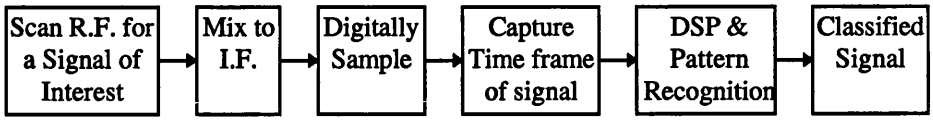


Figure 1.1 : Modulation Recognition Overview

Once the signal is captured digitally there is a great deal of scope to process the signal, and then classification may be performed. There are two broad approaches which may be applied to a pattern recognition problem, which are decision theoretic and non-deterministic methods. The literature has placed more emphasis on non-deterministic techniques, but both forms of pattern recognition are applicable to modulation recognition.

Non Deterministic Techniques

Figure 1.2 describes the techniques involved in the non-deterministic pattern recognition process. Before classification is performed, some parametric estimation may be applied to aid the subsequent pattern recognition process. Parameters which may be estimated are the signal bandwidth, approximate centre frequency and signal to noise ratio.

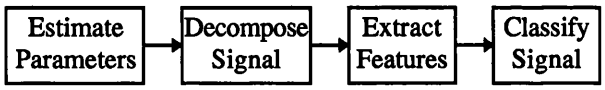


Figure 1.2 : Pattern Recognition Procedure

In modulation recognition, the signal is generally transformed into different forms, which enable a more effective classifier structure and emphasise the information bearing properties of the signal. This stage is termed the decomposition phase. Such transformations may include the envelope, instantaneous frequency and phase of the signal with respect to time. The vector transformations of each stage are described in figure 1.3, where it is seen that the one dimensional signal is transformed into d time dependent signals, which are then processed into n feature vectors to produce a decision of one in m candidate classes.

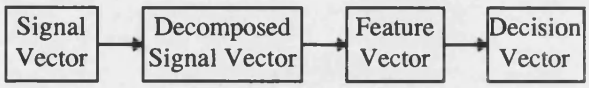


Figure 1.3 : Vector Spaces in the Pattern Recognition Procedure

After the decomposition stage, the feature extraction is performed where the *d* decomposed signals are used to produce time invariant quantities as a descriptive measure of the signal characteristics. These quantities generally number between four and ten [4][27][29] and are known as features. The choice of features in non-deterministic pattern recognition is generally an intuitive exercise, and a number of these are discussed in the following literature review.

The features transform the decomposed signal into an *n* dimensional vector space, and a good choice of feature transformations will place the incoming signal into a set of separable clusters for each signal class in the feature space. A descriptive example is given in figure 1.4 for a transformation into a three dimensional feature space, and clusters of features are illustrated for each modulation class. Generally the signals will not form regular and separable clusters, but will be distributed with different densities about the cluster centre.

Noise has the effect of spreading the features, thus making the clusters less separable and causes each signal class to look more similar. The different features have different levels of robustness to noise, and this is an aspect to be considered in the choice of feature.

Another consideration with the choice of features is that different signals within the same class will often have different parameters e.g. bandwidth, bit rate, frequency deviation and modulation index, and the features must be robust to these variations.

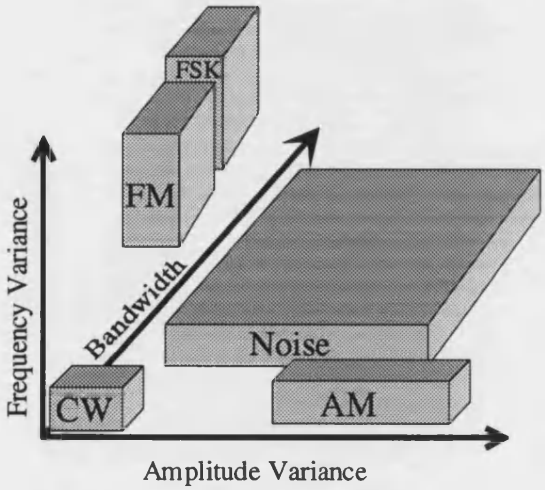


Figure 1.4 : Conceptual Example of three dimensional Feature Space

From the *n* features, classification may be performed, and this is generally achieved through a decision function. The decision function effectively casts boundaries in feature space for each of the *m* classes and the signal is classified depending upon which boundary the feature vector lies within.

In order to determine the boundaries a set of example data is provided to the classifier where an algorithm adapts the decision function boundaries according to an optimisation algorithm. This is known as the training phase. Choosing a 'typical' set of training data is important and choosing the number of examples requires care, as the classifier may be over-trained.

An example of a decision function is given in figure 1.5. This example is a multi-layer perceptron back propagation neural network, which has received much attention since the 1990's [54]. The circles represent 'nodes' which are non-linear functions with multiple inputs that are summed together. The non-linear functions have certain mathematical properties which are required to implement the training algorithm. The connecting lines are gain terms which link the nodes in a feed-forward fashion.

The number of layers of nodes and the number of nodes in each layer are all variables which are often adjusted through trial and error to improve the performance. However, the processing requirements of the network training can be a problem as the number of nodes is increased. There are also problems where the training algorithm converges to local non optimal maxima.

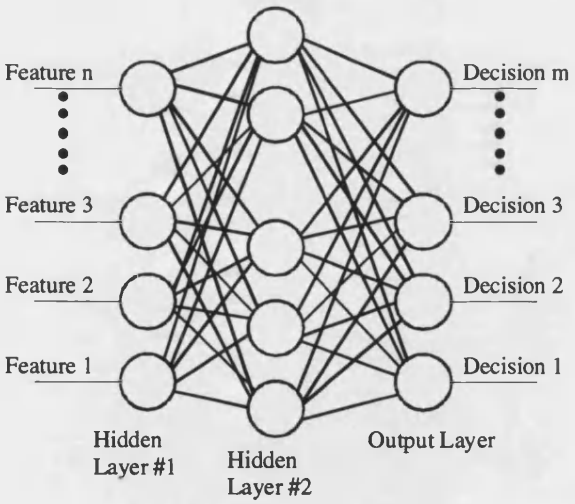


Figure 1.5 : MLP Neural Network Decision Function

The overall effect of the trained network is to produce a highly non-linear function. The outputs are generally trained towards different nodal extremes, which are threshold limited to

produce a binary decision. This action then has the effect of casting boundaries in feature space.

Decision Theoretic Techniques

The preceding discussion was based on non-deterministic methods which assumed no knowledge of the signal characteristics apart from some intuitive transformations. If parametric knowledge about the signal is available, decision theoretic techniques may be applied [54].

Figure 1.6 demonstrates the general procedure of decision theoretic classification. It is common, but not always necessary, to decompose the signal as before. The decomposed signal may be processed further, and an estimate of the signal parameters is made. Examples of such parameters include the signal to noise ratio, generic class of modulation and the amplitude of the signal.

A set of functions are then generated from the parametric data, and each observed decomposed time sample is operated on by the functions. The functions are mathematically derived for each class and represent the probability of the class given the observed data, which is denoted in figure 1.6 by *p*(class *n*| data). The maximum of these then represents the classified signal.

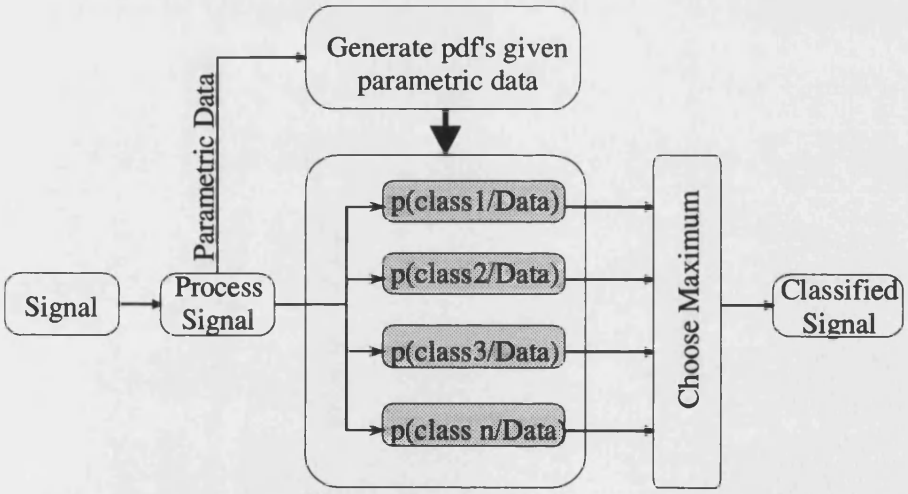


Figure 1.6 : Decision Theoretic Methods for Pattern Recognition

Pattern Recognition Techniques for Modulation Recognition

Modulation recognition is suited to both types of pattern recognition. Digital modulation schemes are generated using mathematical laws, and decision theoretic techniques may be applied for improved performance over non-deterministic methods. In practice the model is

confined within the bounds of a set of assumptions to simplify the problem. Deviations from this paradigm are considered as 'nuisance' terms and will generally act to degrade the system performance.

Analogue signals tend to have short term stationary characteristics, consequently, non-deterministic techniques are generally more appropriate for analogue modulation classification.

In a system which requires the classification of analogue and digital modulation schemes, the pattern recognition may be split into a hierarchical structure as in figure 1.7, where an initial level of pattern recognition is applied to discriminate between analogue and digital modulation schemes and a secondary level is used to classify the individual modulation schemes. It may be feasible that further divisions are made to simplify the implementation.

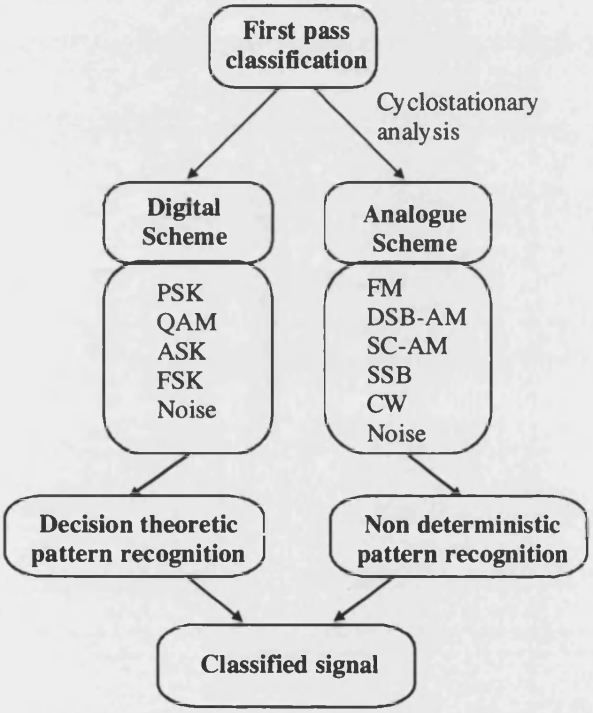


Figure 1.7 : Separation of the Problem

The digital signals possess cyclostationary statistics and under certain transformations the signals contain strong periodic components which may be detected [8][58]. Such periodic components are generally exploited in carrier and symbol synchronisation, and some of these characteristics are discussed in Appendix I.

1.3 Literature Review

The majority of the published literature has been presented from 1985 onwards, with the foundation papers from Leifdtke [25], Callaghan [24] and Jondral [27].

During the period in which the thesis was undertaken a large number of publications have arisen (approximately twenty out of fifty-five), and a discussion of these is included within this chapter. However, the work presented within this thesis is distinct from the other techniques, which is in part due to the broad scope of the problem.

Most of the published material is in the form of conference papers, and there are relatively few which are published in journals. Only three papers appear in IEEE transactions in communications [38][36][39], and all of these papers address a mathematical treatment of decision theoretic techniques. One appears in the IEE transactions [41], and several appear in the Signal Processing Journal [4][6][9][11][27][25][42]. The remaining published papers appear in various conferences, with a large number (approximately sixteen) in the IEEE MILCOM conferences.

The general technique used in the literature is to acquire the received signal through digital sampling. This digitally sampled signal is then applied to the pattern recognition techniques for analysis and classification.

The following review of the literature discusses the decision theoretic and non-deterministic methods in turn. The non-deterministic techniques are discussed first, encompassing the majority of the published literature. This is then followed by the decision theoretic techniques.

1.3.1 Non-deterministic Techniques

It has been identified that the non-deterministic techniques for pattern recognition tend to adopt signal decomposition, feature extraction and classification stages to achieve the task. These processes are discussed in this section within the context of automatic modulation recognition, with reference to the published literature. Also included is the scope of the modulation schemes covered within the literature, along with signal preconditioning and dimensionality reduction techniques

1.3.1.1 Range of Modulation types Considered

Table 1.1 provides a summary of the different modulation types which are classified within the literature. From this table it is seen that the main modulation schemes which are classified are 2ASK, 4ASK, BPSK, QPSK, 8PSK, MSK, 2FSK (Slow), 4FSK (Slow), QAM, SSB,

SC-AM, DSB-AM, FM, CW and Noise. These represent some of the most common modulation schemes which are used across the radio frequencies.

It is also seen from table 1.1 that there is a large variation in the number and type of schemes which each published technique employs. Each scheme added to the classifier increases the dimensionality of the problem, and requires further classifier discriminating powers in order to maintain the classification performance, thus making the problem more difficult to solve effectively. From these wide ranging techniques it is very difficult to provide a useful comparison between the different methods.

CW	[20] [25] [24] [33] [23] [31] [28] [22] [19] [18] [17] [14] [9] [7] [6] [12] [29] [21] [41]
NOISE	[25] [24] [28] [22] [18] [9] [7] [6] [13] [12] [27]
DSB-AM	[24] [28] [18] [4] [14] [9] [10] [7] [6] [13] [12] [11] [51] [26]
SC-AM	[23] [28] [22] [18] [4] [17] [9] [10] [7] [6] [13] [12] [11] [27] [51]
SSB	[24] [33] [23] [28] [22] [18] [4] [14] [7] [6] [13] [12] [11] [27] [51]
FM	[24] [23] [28] [22] [18] [4] [17] [14] [9] [10] [7] [6] [12] [11] [51] [26]
VSB	[4]
FM-AM	[4]
2ASK	[8] [25] [24] [33] [23] [32] [22] [15] [18] [17] [16] [9] [7] [6] [29] [27] [51] [26]
4ASK	[16] [7] [6]
BPSK	[8] [20] [25] [34] [33] [23] [32] [31] [30] [19] [15] [18] [16] [7] [6] [29] [21] [27] [41] [51]
QPSK	[8] [20] [25] [34] [33] [23] [31] [30] [3] [19] [16] [7] [12] [29] [21] [41]
8PSK	[25] [34] [30] [41]
OQPSK	[34] [52]
16QAM	[34] [12]
64QAM	[34]
2FSK	[20] [25] [24] [34] [23] [32] [30] [22] [15] [18] [17] [16] [9] [7] [6] [13] [29] [21] [27] [41] [51] [26]
4FSK	[20] [34] [30] [16] [7] [6] [21] [27] [41]
8FSK	[34] [30] [41]
MSK	[8] [33] [17] [3] [52] [25] [34] [23]

Table 1.1 : A Breakdown of Modulation Schemes Classified Using Non-deterministic Methods

1.3.1.2 Techniques used for Signal Decomposition

The modulation schemes generally have the information bearing signal placed on the envelope or phase of the transmission carrier. The three most common signals used for analysis are the envelope, phase and instantaneous frequency signals. It should be noted that the instantaneous frequency and phase signals have an obvious dependency upon each other. The envelope signal will contain good characterising information for AM, ASK and QAM type signals. The

phase relates in the same way to QAM and PSK type signals, and the frequency signal relates well to FM and FSK type signals.

Of the three mentioned parameters, the envelope is the most simple to derive. In [24] and [22] the envelope is obtained through analogue AM detector circuits, and is then digitally sampled for the subsequent pattern recognition. In most of the other techniques the incoming signal is digitally captured, and then converted into an analytic form. From this analytic representation the envelope is determined by the magnitude of the analytic vector [34] [15] [18] [16] [14] [10] [13] [12] [11] [29] [27] [51].

Using this analytic representation the phase may be deduced by the argument of the analytic vector [25] [19] [15] [18] [3] [16] [12] [27] [51]. The phase obtained through this method is, however, subjected to a modulo 2π transformation. This can be an undesirable effect as it introduces phase discontinuities when the phase moves outside the $(-\pi, \pi]$ boundary.

The most difficult of the three parameters to estimate is the instantaneous frequency. This is because the instantaneous frequency is derived from the derivative of phase, and can at best only be an approximation to the true instantaneous frequency. The most common method of obtaining the instantaneous frequency is through a first order differential approximation to the phase [25] [34] [15] [18] [16] [14] [6] [12] [27]. Another common method is achieved by differentiating the phase in terms of the quadrature channels [28] [19] [10] [13] which provides a computationally simple solution. The zero crossing rate has been applied as an estimator [24] [7] [6] [41] [51] [59], which is one of the most simple estimators which can be applied. Two slightly more unusual techniques which have been applied are an autoregressive technique [20] [21] which is used to search for the peak of the spectrum, and a method based upon an approximation to the first order differential of phase [29].

The transformations of envelope frequency and phase provide a time domain transformation of the signal. Other transformations may be applied in the frequency domain and the time frequency domain.

The technique presented in [23] performs classification based solely on the frequency bins of a Welch averaged periodogram. Frequency domain parameters can be useful for averaging time dependent aspects of the signal, and they can also highlight features which are very difficult to isolate in the time domain. However, care must be taken when using spectral techniques, as classification in [23] is dependent upon the bandwidth of the signal, and may not prove to be a robust estimator.

Two time-frequency techniques have been discussed, describing the use of the Wigner distribution for modulation classification [1] [2]. In [1] a brief discussion is provided to point out the pseudo-Wigner tool as a means for signal translation in modulation recognition. The continuous version of the Wigner distribution is discussed in [2], and a set of patterns is described for different modulation schemes, but no classification technique is applied. It should be noted that the generation of a continuous Wigner distribution is currently not a practical method to implement, and will probably rely on acousto-optical techniques in the future.

One problem with time frequency techniques is that the dimension of the pattern recognition process becomes extremely large, and problems of bandwidth scaling still exist. A problem with discrete time-frequency techniques is that unwanted effects occur due to a discrete time approximation.

The wavelet transform has been investigated in [30] [31] [32], and a different approach has been adopted in each case. In [30] a very simple wavelet function was used, and the wavelet transformation was simply due to the difference between a sum of seven samples on each side of a symbol.

In [32] the wavelet transform is achieved through a wavelet packet approach, which is similar to the use of sub-band decomposition with additional stages. The energy from each decomposition is used as a feature in a neural network to discriminate between PSK, ASK and FSK modulation schemes.

An interesting technique is presented in [31] where classification of PSK signals is based upon the number of observed symbol transitions. Assuming that symbols are equiprobable, higher order PSK types will have a higher probability of different symbol levels between adjacent symbols. In order to detect symbol changes, a continuous wavelet transform with a Morlet wavelet is used.

A spectral method based upon the cyclostationary behaviour of digitally modulated signals is proposed in [8]. This uses a power transformation to transform the signal into one with a discrete spectral component. The component may then be detected, and depending upon the power law which was applied, the signal may be classified. This is similar to a classical carrier detection technique, and provides good discriminating capabilities in noise.

A interesting approach for the classification between OQPSK and MSK has been taken in [52], where singular value decomposition (SVD) is applied to the signal in order to exploit the

features of MSK. It is found that one of the SVD vectors is a useful discriminating feature between the two modulation schemes. This feature is illustrated in an ideal noiseless system where the vector is zero for OQPSK and a linearly increasing line for MSK. This noiseless MSK vector is correlated with a vector generated from the signal under test, and the result is compared against a threshold.

1.3.1.3 Signal Preconditioning

Part of the signal decomposition requires some preconditioning of the signal, and one very difficult requirement in most techniques is that the carrier frequency is removed from the signal.

In any technique performing the classification of PSK and QAM signals, it is essential to have an accurate knowledge of the carrier frequency of the signal, as a carrier offset causes the phase signal to become distorted. In some techniques it is assumed that the carrier is accurately known [25] [34] [30] [15] [4] [16] [7] [52], and in others techniques are described for carrier frequency estimation.

A digital phase locked loop is used in [27] to highlight a carrier frequency reference of BPSK signals. A similar method is applied in [18] except that the phase locked loop is replaced with a spectral search followed by a bandpass filter. In [41] [6] [19] [20] [21] [5] [12] the carrier frequency is estimated as the mean of the instantaneous frequency with the frequency spikes removed, and in [29] estimation is taken from the peak of an instantaneous frequency histogram.

The phase becomes meaningless when the signal level is zero and, in general, the phase is not useful when the envelope of the signal is small, as the effects of noise will dominate the phase. In [4] and [28] the phase samples are omitted when the envelope drops below a threshold. In [29] a similar technique is used, but the threshold is set by a fuzzy algorithm.

1.3.1.4 Feature Extraction

The feature extraction stage is the most important stage within the pattern recognition process, and many techniques are found within the literature. The aim of a feature extraction technique in traditional pattern recognition methods is to produce a quantitative value for the characteristics of the signal. A good choice of feature vector will enable a better separation of clusters within feature space for classification. The effects of noise will cause the feature clusters for each class to tend to merge together, and a good feature will be capable of suppressing the effects of noise in a relative capacity.

The choice of features is generally an intuitive exercise, and a wide variety of techniques have been proposed. The most popular features are the moments of the decomposed signals. The variance of envelope is used in [25] [24] [28] [18] [16] [12] [3] [51] [26] [34] [13], but in all cases the signal energy is normalised in order to produce a result which is independent of the received signal strength. This feature is useful for discriminating amplitude modulated signals from constant envelope signals. In [11] the normalised envelope variance was analysed with respect to an uncorrelated Gaussian speech modulation, and was used to discriminate between different analogue amplitude modulated signals.

The frequency variance feature is employed in [20] [25] [24] [22] [14] [12] [21] [41] [51] [16] [26] [34] [13] in order to highlight the frequency modulated signals. The phase variance is used in [18] [4] [16] [12] [51] [34], but it should be noted that there will be overlap between the discriminating properties of the phase and frequency variance.

The third and fourth central moments (skewness and kurtosis) of envelope frequency and phase are applied to the methods in [34] [16] [12] [13, not phase]. These moments partly describe the shape of the probability density functions. In [16] the mean and median of the three signals are also used.

A set of features may be expressed through the histogram of envelope frequency and phase, which describes the probability density function of the modulation parameters, and can be argued to contain all the information of the corresponding central moments. The modes of digitally modulated signals are parameters which are emphasised through the histogram. This technique has been applied to [27] [6] [15] where the histogram bins are fed directly into the classifier. Other techniques perform a further process, which fringes upon that of classification [30] [7] [29] [41] [26], where the number of signal modes are determined. This reduces the dimensionality of the features, and there is more control over the method of mode determination.

In order to remove the requirement of a zero phase reference, a histogram of the phase difference between two symbols is applied in [25] [41] [19] where knowledge of the symbol rate and timing is required. A template is then used with [25] [41] to classify the number of modes, and a threshold is applied to [19]. This is a useful technique when classifying PSK signals.

In order to combat the effects of bias due to noise, a set of features were developed in [9] which were asymptotically unbiased in the presence of Gaussian noise. Three features were

developed, and preliminary results were discussed in [9] and [10]. The features were extended further in [13], where the second, third and fourth central moments were applied.

Spectral line detection has been used for feature detection in [8] [18] and [3]. In [8] a power law applied to an analytic signal is used to discriminate ASK, PSK, FSK and MSK modulated signals. In [18] the variance of signal spectrum is analysed through the signal and it's square with components removed below a threshold. This essentially determines whether there is a discrete carrier component present. In [3] a power law is applied to the envelope of the signal following bandpass filtering in order to discriminate between MSK and PSK signals. The spectral line in this case is due to the cyclostationary behaviour of the envelope through the symbol transitions, and is prominent in PSK. In all three cases the FFT is used for frequency analysis.

Bandwidth information is used in [20] [21] where a short time analysis of the bandwidth is determined using autoregressive techniques. This is used to discriminate between analogue and digitally modulated signals.

1.3.1.5 Dimensional Reduction

In pattern recognition it can be tempting to apply a large number of features to a classifier in order to provide the classifier with as much information as possible. However, as the number of features is increased, the complexity of the classifier and the classifier training time will also increase. By introducing too many features it is usually found that some features may have no discriminating powers and other features may have a large amount of overlap.

In order to reduce the dimension and redundancy of the features, a technique known as Principal Component Analysis (PCA) may be applied. This is also known as the Karhunen-Loève transform, which essentially performs a linear transformation on the feature vector to a new set of vectors, and the discriminating capabilities of each output from the transform may be quantified. The lower discriminating outputs may be identified and then omitted, thus reducing the dimension of the input. This technique has been performed in [15] and [27]. In [27] the features were reduced from 192 histogram bins to 93 features.

1.3.1.6 Classification Techniques

Classification is the final part of the pattern recognition process, and may be performed using a number of techniques. Two common techniques for classification are those employing decision functions and those classified through the distance from a class. Included in the

decision function form of classification are neural networks which utilise highly non-linear functions.

The most simple decision function is the linear classifier, where hyper-planes are used to cast boundaries across classes. The placement of the planes may be performed using different algorithms, and this technique has been adopted in [17] [26] [6] [13] [27]. A more complicated arrangement derived from hyper-quadratics was implemented in [27] [18] [33], and in [27] a performance comparison is provided between the linear and quadratic classifiers.

Neural networks have been investigated in [34] [32] [23] [16] [12] [21], and in most cases a multi layer perceptron network is used. In [21] and [34] the complexity of the network is reduced by splitting the problem up using a hierarchical approach. The minimum distance classifier has not been a popular technique and is only suggested as an application in [24].

Most of the other classification techniques have been achieved through user determined thresholds on decision trees [19] [4] [14] [7] [29] [51] [20] [25] [8] [3] [30] [11]. This technique can enable flexibility in the classification process and the mixture of discrete features such as the number of modes with analogue features. In [14] [11] and [8] some analysis and theoretical reasoning has been applied to the thresholds, and in [41] a decision tree has been used in conjunction with some maximum likelihood methods.

1.3.1.7 Measurement Conditions

In most of the published literature the results are from simulations, the noise is Gaussian, no fading (frequency selective or Rayleigh) is included, non-linearity and bandlimiting effects are neglected and interference is assumed to be absent.

There are a small number of exceptions. [25] performs simulations of adjacent channel interference, carrier offset and symbol timing error, [33] and [27] use real signals stored on magnetic media, [8] generates digital signals using Gaussian pulse shapes, [34] applies filtering to the signal and [28] analyses the effect of non-Gaussian noise.

Most of the other published work assumes that an ideal signal is present in the presence of Gaussian white noise.

1.3.2 Decision Theoretic Techniques

The decision theoretic methods are based upon mathematical models, and a detailed description is difficult to achieve without a discussion of the mathematics. This section briefly

discusses the development of the decision theoretic techniques in modulation recognition and a mathematical treatment of some of the methods is provided within chapter 3 of the thesis.

A number of decision theoretic techniques are found in the published literature and these have been applied to the classification of digitally modulated signals [44] [45] [46] [47] [38] [39] [53] [35] [37] [36] [41] [42] [43] [40]. Most of these papers have been authored or co-authored by A. Polydoros from the University of Southern California [44] [45] [46] [47] [38] [39] [53] and S. Soliman of Qualcomm [35] [37] [36] [41] [42] [43].

One final paper [40] is an extension of [36] which approximates the pdf of phase using a Tikhonov distribution instead of an exact distribution. The paper looks into a transformation of the phase signal through a cosine, and determines an exact distribution. However the authors of [36] provided an exact distribution without a transformation in [35] and [42].

Most of the papers have been applied to the classification of the number of levels on PSK signals [35] [37] [36] [41] [42] [38] [39] [40]. Two papers discuss the classification of the index pattern of continuous phase modulation [45] [47] and the rest are applied to QAM/PSK classification [44] [46] [53].

All of the decision theoretic techniques discussed assume that the signal is in AGWN and that the carrier frequency and symbol timing information is available.

The first paper to introduce PSK classification was [38] where a set of approximations are applied to a likelihood function for the discrimination of BPSK and QPSK. The classifier likelihood function is comprised of the in-phase and quadrature signals squared and a set of cross terms. This is then compared with a threshold for classification. The threshold for statistical comparison is not determined analytically, but instead the conditional pdf of the two classes is derived analytically. The threshold is presumed to be determined numerically by determining the intersection point of the two pdf's. Following this initial paper [44] [37] [36] were published in 1991/1992.

In [44] further work is provided along the lines of [38], and a more general likelihood function is presented for PSK and QAM signals at low SNR. An approximation to the likelihood function is proposed using an M^{th} law method and a threshold is analytically determined for discriminating between two PSK signals. Discrimination between PSK and QAM signals is also suggested using the M^{th} law, but only certain constellations can be easily discriminated and there are no analytical methods described for determining the necessary decision thresholds.

Two techniques based upon the phase samples instead of the quadrature samples are proposed in [36] and [37], which have the advantage of not requiring knowledge of the signal amplitude. Both techniques incorporate the flexibility to enable classification of a number of modulation schemes as opposed to only two in [44] and [38].

In [36] classification is based upon the statistical moments of the phase signal. The statistical moment of the phase is then compared to a set of thresholds for classification. In this paper the pdf of phase is approximated by a Tikhonov distribution in order to produce an analytical derivation of the thresholds. The Tikhonov approximation is replaced by an exact representation in [35] and [42].

Another phase based classifier is developed in [37] for the classification of PSK signals, and used maximum likelihood techniques for classification. The results from this paper enable better discriminating performance than those of [38] and [36], but the computational burden is high.

A recent paper [39] has described the derivation of some of the results in [44] in greater detail. This looks into the classification of PSK signals using an M^{th} power law, and provides an extension which enables a number of PSK schemes to be classified instead of the binary example in [38] and [44].

An extension of the QAM classification techniques of [44] is discussed in [46], and it is found that methods for deriving thresholds are very difficult and a single term approximation is not always suitable. Some methods are applied here to model the distributions, determine thresholds and introduce additional terms in the approximation.

1.3.3 Summary of the Literature Search

The results of the literature search have highlighted the fact that there are many different approaches applied to modulation recognition. It has been seen that most of the non-deterministic techniques attempt to classify different modulation schemes, and will generally apply different techniques for feature extraction. The results are often difficult to compare, and in a number of reported techniques, information has been omitted from the paper which makes the implementation of the published technique difficult.

There are very few systems which have been implemented in a commercial product, and most of the published results are based upon system simulations, which makes the subject somewhat theoretical. However, there are still many techniques to explore before considerable effort is well spent in characterising the techniques under a large range of conditions.

Most systems in the published literature digitally sample an analysis time frame and decompose the signal into the envelope, phase and instantaneous frequency from an analytic signal. It has been found that the generation of an analytic signal and the estimation of the instantaneous frequency has been achieved through a number of techniques, and there is very little justification provided for taking a particular approach.

The decision theoretic structures have been developed for digitally modulated signals, and have mainly been tackled in the published literature by two groups, and the majority of the literature is based on the classification of PSK signals.

1.4 Aspects Tackled by the Thesis

From the literature survey it has become apparent that there are a large number of classifiers implementing different techniques on an ad-hoc basis. The work undertaken in this thesis does not attempt to provide another such classifier, but attempts to isolate specific issues.

A) Signal Modelling

In this thesis, new signal models are developed, and the work is intended as a tool in the analysis of modulation recognition techniques. The main areas considered are the pdf and related statistics for a sinusoid in AGWN of the three primary decomposed signals in modulation recognition : the envelope, phase and instantaneous frequency.

A number of instantaneous frequency estimation techniques, including those in the published literature, are compared in terms of performance and complexity. An appropriate instantaneous frequency estimator is then selected for modulation recognition purposes, where the pdf and related statistics may then be modelled.

The work provides new results and a useful summary of existing results, and finds some direct use within the decision theoretic classification considered within this thesis.

B) PSK Classification

A number of decision theoretic PSK classification techniques have been developed mainly by two groups in the published literature. The various techniques require different a-priori information and have different characteristics, but they all assume that the carrier frequency is accurately known.

A number of novel techniques are developed for classifying PSK signals using decision theoretic methods. These are compared with existing techniques and have been found to have

different characteristics and parametric knowledge requirements to those in the published literature.

Some of the developed techniques also assume that the carrier frequency is accurately known, but also a new set of decision theoretic techniques are developed which do not require accurate carrier frequency knowledge.

The published literature has generally provided a discussion of classifier performance under ideal conditions, and work is presented to analyse the classifier performance of the various techniques when non-ideal effects are introduced.

C) Other Aspects

The generation of an analytic signal is fundamental in the decomposition of the signal, and a number of techniques have been proposed in the published literature. Work has been undertaken to investigate four different techniques, and provides a direct comparison between each of them.

1.5 Thesis Organisation

The structure of the thesis is organised as follows :

Chapter 2 derives and outlines various mathematical models for different signals in the presence of AGWN. The results from the work in this chapter are applied to the decision theoretic methods for PSK classification in chapter 3. The primary results are summarised in the main body, and the derivations are provided in appendices held within the chapter. Also included in this chapter is a discussion of an appropriate digital instantaneous frequency estimator for modulation recognition applications, and the statistics are derived for this choice of estimator.

Chapter 3 discusses new decision theoretic techniques for the classification of the number of levels on PSK signals. The work derives new classifier structures and also summarises and improves aspects of some existing structures.

Chapter 4 provides a comparison of the different techniques in terms of classification performance when certain assumptions in the decision theoretic paradigm are broken down.

Conclusions of the work into decision theoretic PSK classification are presented in Chapter 5 with an outline of key areas for future work. Also included within this chapter are some more general observations for future research into modulation recognition as a whole.

Appendix I discusses the problems associated with the parametric estimation required for decision theoretic classification, including a description of carrier and symbol synchronisation techniques and methods for estimating the signal SNR. Appendix II discusses different DSP methods for the generation of an analytic signal, and the associated performance of each technique.

Finally, a list of the publications which have arisen from the work conducted within this thesis are provided in Appendix III.

1.6 References

[1] S.C Pei, T.Y Wang, "Modulation Signal Classification by the Wigner Distribution", ISSPA 87, pp. 269-274

[2] B. Sankar, N. Tugbay, "The use of the Wigner-Ville Distribution (WVD) for the classification and the Analysis of Modulated Techniques", ISCIS 88, pp. 75-82

[3] M. Hagiwara, M. Nakagawa, "Estimation of an Input Signal Type", GLOBECOM '87, 7.6.1-7.6.5, pp. 254-258

[4] A.K. Nandi, E.E. Azzouz, "Automatic Analogue Modulation Recognition", Signal Processing, Vol. 46, 1995. pp.211-222

[5] R.J. Inkol, R.H. Saper, "A New Algorithm for Signal Classification", Proc. NAECON '93, pp. 320-326.

[6] L. Vergara Dominguez, J.M. Páez Borrallo, J. Portillo Garcia, "A General Approach to the Automatic Classification of Radiocommunication signals", Signal Processing, Vol. 22, 1991, pp. 239-250.

[7] J. Portillo Garcia, J.P. Sancho-Marco, L. Vergara Dominguez, J.M. Páez Borrallo, Ruiz-Mezcua, "A Microcomputer-Based General Architecture for Radiocommunication Signal Classification and Digital Demodulation", Proc. EUSIPCO '90, vol. 3, pp. 1919-1922.

[8] J. Reichert, "Automatic Classification of Communication Signals Using Higher Order Statistics", Proc. ICASSP'92, Vol. 5, pp. V.221-V.224

[9] J. Aisbett, "Automatic Modulation Recognition Using Time Domain Parameters", Signal processing, Vol 13, 1987, pp. 323-328

[10] K.B. Parcell, J. Aisbett, "Automatic Modulation", Signal processing, ISSPA' 87, vol. 2, pp. 871-874

[11] Y.T. Chan, L.G. Gadbois, "Identification of the Modulation Type of a Signal", Signal Processing, Vol. 16, 1989, pp. 149-154

[12] J.E. Whelchel, D.L. McNeill, R.D. Hughes, M.M. Loos, "Signal Understanding : an Artificial Intelligence Approach to Modulation Classification", Proc. IEEE International Workshop on Tools for A.I. 1989, pp. 231-236

[13] G.A. Einicke, "A Flexible Architecture for Real-time Modulation Recognition", ASSPA '89, pp. 133-137.

[14] P.M. Fabrizi, L.B. Lopes, G.B. Lockhart, "Receiver Recognition of Analogue Modulation Type", IERE Radio Receivers and Associated Systems '86, pp. 135-140.

[15] E.R. Adams, P.C.J. Hill, C.N. Kempson, "A Statistical Approach to Modulation Recognition for Communications Signal Monitoring and Surveillance", IERE Sept. 1988, pp. 31-37

[16] P.C.J. Hill, G.R. Orzeszko, "Performance Comparison of Neural Network and Statistical Discriminant Processing Techniques for Automatic Modulation Recognition", Proc. SPIE, Vol. 1469, Applications of Artificial Neural Networks, 1991, pp. 329-340

[17] A. Martin, "A Signal Analysis and Classification Strategy for Implementation in an EW Communications Receiver", IEE Radio Receivers and Associated systems, IEE conf. 325, July 1990, pp. 222-226

[18] J.E. Hipp, "Modulation Classification Based on Statistical Moments", IEEE MILCOM '86, pp. 20.2.1-20.2.6

[19] R.J. Mammone, R.J. Rothaker, C.I. Podilchuk, S. Davidovici, D.L. Schilling, "Estimation of Carrier Frequency, Modulation Type and Bit Rate of a Modulation Signal", IEEE ICC '87, pp. 28.4.1- 28.4.7

[20] K. Assaleh, K.R. Farrell, R.J. Mammone, "A New Method of Modulation Classification for Digitally Modulated Signals", MILCOM '92, pp. 30.5.1-30.5.5

[21] K.R. Farrell, R.J. Mammone, "Modulation Classification Using a Neural Tree Network", MILCOM '93, pp.1028-1032

[22] P.M. Petrovic, Z.B. Krsmanovic, N.K. Remenski, "An Automatic VHF Signal Classifier", MELCON '89, pp. 385-387

[23] N. Ghani, R. Lamontagne, "Neural Networks Applied to the Classification of Spectral Features for Automatic Modulation Recognition", IEEE MILCOM'93, pp. 111-115.

[24] T.G. Callaghan, J.L. Perry, J.K. Tjho, "Sampling and Algorithms aid Modulation Recognition", Microwaves & R.F., September 1985, pp. 117-121

[25] F.F. Liedtke, "Computer Simulation of an Automatic Classification Procedure for Digitally Modulated Communication Signals with Unknown Parameters", Signal Processing, Vol. 6, No. 4, August 1984, pp. 311-323

[26] P.N.A.P. Rao, G. Boopathy, "Analysis and Classification of Communication Signals", Defence Science Journal (India), Vol. 35, No. 3, July 1985, pp. 367-374.

[27] F. Jondral, "Automatic Classification of High Frequency Signals", Signal Processing, Vol. 9, No. 3, October 1985, pp. 177-190

[28] P.A.J. Nagy , "Analysis of a Method for Classification of Analogue Modulated Radio Signals", EUSIPCO '94

[29] P.A.J. Nagy, "A Modulation Classifier for Multi Channel Systems and Multi Transmitter Situations", IEEE MILCOM '94, Vol. 3, pp. 816-820

[30] K.C. Ho, W. Prokopiw, Y.T. Chan, "Modulation Identification by the Wavelet Transform", MILCOM '95, pp. 886-890

[31] Yu-Chuan Lin, C.-C Jay Kuo, , "Modulation Classification using Wavelet Transform", Proc. SPIE, Vol. 2303, 1994, pp. 260-271

[32] N. P. Ta, "A Wavelet Packet Approach to Radio Signal Classification", Proc. IEEE Symposium in Time Frequency Time Scale Analysis, 1994, pp. 508-511

[33] G. Feyh, M. Kükenwaitz, J. Reichert, "HF-Signal Surveillance : Signal Detection, Classification and Parameter Estimation", MILCOM'94, pp. 755-759

[34] C. Louis, P. Sehier, "Automatic Modulation Recognition with a Hierarchical Neural Network", MILCOM '94, pp. 713-717

[35] Y. Yang and S.S. Soliman, "Statistical Moments Based Classifier For MPSK Signals", Proc. IEEE GLOBECOM, pp. 2.7.1-2.7.5, 1991

[36] S.S. Soliman, S.Z. Hsue, "Signal Classification Using Statistical Moments", IEEE trans. Communications, Vol. 40, No. 5, pp. 908-916, May 1992

[37] Y. Yang, S.S. Soliman, "Optimum Classifier For M-ary PSK Signals", Proc. IEEE International Conference On Communications, pp. 52.3.1-52.3.5, 1991

[38] A. Polydoros, K.K. Kim, "On The Detection And Classification Of Quadrature Digital Modulations In Broad Band Noise", IEEE Trans. Communications, Vol.38, No.8, pp. 431-436, August 1990

[39] C.Y. Huang, A. Polydoros, "Likelihood Methods For MPSK Modulation Classification, IEEE Trans. Communications", Vol. 43, No. 2/3/4, pp. 1493-1504, Feb/March/April 1995

[40] Y.H. Han, W.C. Lee, D.W. Yoon, "On the Error Rate Evaluation and Modulation Classification for Coherent and Noncoherent PSK Signals using the Transformation of Random Variable", ICC'93, pp. 1508-1514

[41] S.Z. Hsue, S.S. Soliman, "Automatic Modulation Classification Using Zero Crossing", IEE Proceedings, Vo. 137, Pt. F, No. 6, Dec. 1990, pp. 459-464

[42] Y. Yang, S.S. Soliman, "An Improved Moment-Based Algorithm for Signal Classification", Signal Processing, Vol. 43, No. 3, 1995, pp. 231-244

[43] S.Z. Hsue, S.S. Soliman, "Automatic Modulation Recognition of Digitally Modulated Signals", IEEE MILCOM'89, pp. 645-649

[44] C.Y. Hwang, A. Polydoros, "Advanced Methods for Digital Quadrature and Offset Modulation Classification", IEEE MILCOM '91, pp. 841-845

[45] C.D. Chung, A. Polydoros, "Envelope Based Schemes for Continuous-Phase Binary Frequency-Shift-Keyed Modulations", IEEE MILCOM'94, pp. 796-800

[46] C.S. Long, K.M. Chugg, A. Polydoros, "Further Results in Likelihood Classification of QAM Signals", IEEE MILCOM'94, pp. 57-61

[47] C.Y. Huang, A. Polydoros, "Two Small-SNR Classification Rules for CPM", IEEE MILCOM'92, pp. 1236-1240

[48] R.F. Schneider, D.C. Chu, "Modulation Recognition of Spread Spectrum Signals Using Modulation Domain Measurements", IEEE MILCOM'91, pp. 819-826

[49] F. Jondral, "Principles of Automatic Signal Classification", Proc. IERE, No. 68, July '86, pp. 147-153

[50] F. Jondral, "Foundations of Automatic Modulation Classification", Proc. Stochastic Models & Methods in Information Processing, pp. 201-206, 1989

[51] S.B. McMillan, B.P Flanagan, T.K. Doong, "Determination of the Modulation Type of Communication Signals", IEEE ICASSP'90, pp. 1683-1686

[52] N.M. Marinovich, N.M., Nelson, D., Cohen, L., Umesh, S., "Classification of Digital Modulation Types", Proc. SPIE, Vol. 2563, July 1995, pp. 125-143

[53] Y.C. Lin, C.C. Jay Kuo, "Modulation Recognition via Sequential Probability Ratio Test", Proc. SPIE, Vol. 2563, July 1995, pp. 382-393

[54] Y. Pao, "Adaptive Pattern Recognition and Neural Networks", Addison-Wesley, 1989

[55] P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Classification of PSK Signals using the DFT of Phase Histogram", ICASSP-95, Vol. 3, pp. 1868-1871, May 1995

[56] P.C. Sapiano, J.D. Martin, , "Maximum Likelihood PSK Classification using the DFT of Phase Histogram", GLOBECOM '95, vol. 2, pp 1029-1033

[57] P.C. Sapiano, J.D. Martin, "Identification of PSK Signals", Proc. IEE Radio Receivers and Associated Systems Conference, September 1995, No. 415, pp. 95-99

[58] W.A. Gardner, "Statistical Spectral Analysis : a Nonprobabilistic theory", Prentice Hall, N.J., 1988

[59] B. Kedem, "Spectral Analysis and Discrimination by Zero-Crossings", Proc. IEEE, Vol. 74, No. 11, Nov. 1986, pp. 1477-1493

Chapter 2 : Signal Models

2. Signal Models

2.1 Outline

Modulation classification is a pattern recognition problem, and it is important to have knowledge of the signal characteristics. The analysis of various aspects of the signal is therefore an essential component of modulation recognition.

These signal models may only be analysed within a limited range of conditions, which are generally confined to a signal in Additive Gaussian White Noise (AGWN). The models will attempt to describe the pdf and its associated moments for a particular signal components.

The aim of this section of the thesis is to summarise the various models which exist, and develop additional models. These models may then be applied to decision theoretic classification, or may be used to enhance the understanding of non-deterministic techniques.

The main attributes which are analysed are the pdf of envelope, phase and instantaneous frequency in AGWN for different signals. The instantaneous frequency may only be estimated using approximations, and a number of different techniques may be applied in the estimation. Section 2.3.6 compares a number of different instantaneous frequency models and concludes upon a suitable estimator for automatic modulation recognition applications. The statistical properties and pdf of the estimator are then discussed.

The specific areas which have been progressed by this research are the modelling of the pdf of phase for a sinusoid and PSK signals, and the pdf of instantaneous frequency for a sinusoid in Gaussian noise, which is reflected in the published papers [1] and [2]. Also, new numerical methods have been developed for the generation of various pdfs which provide efficient computational evaluation. These results have a direct impact on the system efficiency at the classification stage.

This chapter contains a number of Appendices describing the mathematical detail and derivations, with the main body of the chapter summarising the results. This is expected to provide a clearer picture of the work and will enable the reader easily to pick out results.

2.2 Assumptions

The models which are developed in this section assume a signal in AGWN. When multi-level PSK is examined, it is assumed that there are an equal number of occurrences of each symbol,

and that the envelope is constant, which is a reasonable assumption with a system including a matched filter, but deteriorates in a bandlimited system without matched filtering.

An analytic signal is assumed present for the analysis of the envelope, phase and instantaneous frequency of the signal. Techniques for generating a digital analytic signal are described in Appendix II, where different practical aspects are discussed.

2.3 Development

The sinusoid is the starting point for the development of a number of different cases, and the pdf of envelope and phase of a sinusoid in AGWN has been investigated in some detail in the past [3].

The joint pdf of envelope z and phase ϕ for a sinusoid of amplitude A in AGWN of variance σ^2 is given by [3] :

$$p(z,\phi)=\frac{z}{2\pi\sigma^2}\exp\left[\frac{z^2+A^2-2Az\cos(\phi)}{-2\sigma^2}\right] \tag{2.1}$$

2.3.1 PDF of Envelope

The pdf of phase or envelope may be expressed independently by eliminating the opposing variable. This may be achieved by integrating the expressions across the full range with respect to the opposing variable. For the pdf of envelope this yields :

$$p(z)=\frac{z}{2\pi\sigma^2}\exp\left[\frac{z^2+A^2}{-2\sigma^2}\right]I_0\left(\frac{Az}{\sigma^2}\right) \tag{2.2}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind. The moments of the envelope are given by [3] :

$$E\{z^n\}=(2\sigma^2)^{\frac{n}{2}}\Gamma\left(\tfrac{n}{2}+1\right){}_1F_1\left(-\tfrac{n}{2};1;-\rho\right) \tag{2.3}$$

where ${}_1F_1(a;b;z)$ is the confluent geometric function [4] given by :

$${}_1F_1(a;b;z)=1+\frac{a(a+1)}{b(b+1)}\frac{z^2}{2!}+\cdots\frac{a(a+1)\cdots(a+n-1)}{b(b+1)\cdots(b+n-1)}\frac{z^n}{n!} \tag{2.4}$$

using [5] the following moments for the envelope may be deduced as :

n	Moment
1	$\sigma\sqrt{\frac{\pi}{2}}e^{-\frac{\rho}{2}}\left[(1+\rho)I_0\left(\frac{\rho}{2}\right)+\rho I_1\left(\frac{\rho}{2}\right)\right]$
2	$2\sigma^2(\rho+1)$
4	$8\sigma^4\left(2\rho^4+4\rho^2+1\right)$
6	$48\sigma^6\left(\frac{4}{3}\rho^6+12\rho^4+3\rho^2+1\right)$

Table 2.1 : Moments Of The Envelope of a Sinusoid in Gaussian Noise

2.3.2 PDF of Phase

By integrating (2.1) with respect to the envelope, in the interval (0,∞), the pdf of phase for a sinusoid in AGWN is given by [3] :

$$f(\phi)=\frac{1}{2\pi}e^{-\rho}+\frac{1}{2}\sqrt{\frac{\rho}{\pi}}\cos(\phi)e^{-\rho\sin^2(\phi)}\left[1+\operatorname{erf}\left(\sqrt{\rho}\cos(\phi)\right)\right] \tag{2.5}$$

where

$$\operatorname{erf}(x)=\int_0^x e^{-t^2}\,dt \tag{2.6}$$

The pdf of phase given in equation (2.5) is bound in the interval [-π,π). Such a function may be expressed in terms of a Fourier series [6] as :

$$f(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{m=1}^{\infty}b_m\cos(m\phi) \tag{2.7}$$

This form of pdf is less commonly used than (2.5), but it will be found that it yields some very useful results in modulation recognition. The Fourier coefficients are derived from special functions in [7], and from first principles in Appendix 2.A, which is a more detailed description of the account given in [1].

It is found that the Fourier coefficients may be written in three forms ,

Series form :

$$b_m=e^{-\rho}\rho^{\frac{m}{2}}\sum_{k=0}^{\infty}\frac{\Gamma\left(k+\frac{m}{2}+1\right)}{k!\Gamma(k+m+1)}\rho^k \tag{2.8}$$

Hypergeometric form :

$$b_m = \rho^{\frac{m}{2}} \frac{\Gamma\left(\frac{m}{2}+1\right)}{\Gamma(m+1)} {}_1F_1\left(\tfrac{m}{2}; m+1; \rho\right) \tag{2.9}$$

and modified Bessel function form :

$$b_m = \frac{\sqrt{\rho\pi}\, e^{-\frac{\rho}{2}}}{2}\left[I_{\frac{m+1}{2}}\left(\tfrac{\rho}{2}\right) + I_{\frac{m-1}{2}}\left(\tfrac{\rho}{2}\right)\right] \tag{2.10}$$

This final form is expressed in terms of Bessel functions of order integer and integer plus a half. Some properties of the Bessel functions of order integer plus a half are discussed in Appendix 2.B with reference to the Fourier coefficients.

In [9] the coefficients are evaluated using the series form of (2.8). A more useful technique for the evaluation of the function is the modified Bessel function form (2.10), and Appendix 2.C describes an efficient technique for generating the Fourier coefficients, which is also described in brief in [1].

A set of cosine evaluations is also required in order to generate the pdf and an efficient technique for generation which is upwardly stable is provided in Appendix 2.D.

The pdf may be extended to the pdf of phase for *M* level PSK (Appendix 2.E) to give :

$$f_M(\phi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} (-1)^n b_{nM} \cos(nM\phi) \tag{2.11}$$

This result shows that *M*PSK signals have the same form of pdf of phase as that of a sinusoid in AGWN, except that the harmonics are decimated by a factor equal to the PSK type. On closer inspection of (2.11) it is seen that the PSK Fourier harmonics will consist of Bessel functions of order integer plus a half, and evaluation requires only one of the two iterative algorithms detailed in Appendix 2.C.

2.3.3 Low SNR Development

The Fourier series expression for pdf of phase can obviously not be evaluated to an infinite number of terms, and so the series must be truncated.

The graph in figure 2.1 shows how the Fourier harmonic magnitudes vary against SNR for the harmonics 1,2,4 and 8. From the graph it can be seen that the separation between adjacent harmonic components increases as the SNR decreases. This indicates that the pdf may be defined accurately by a finite number of terms.

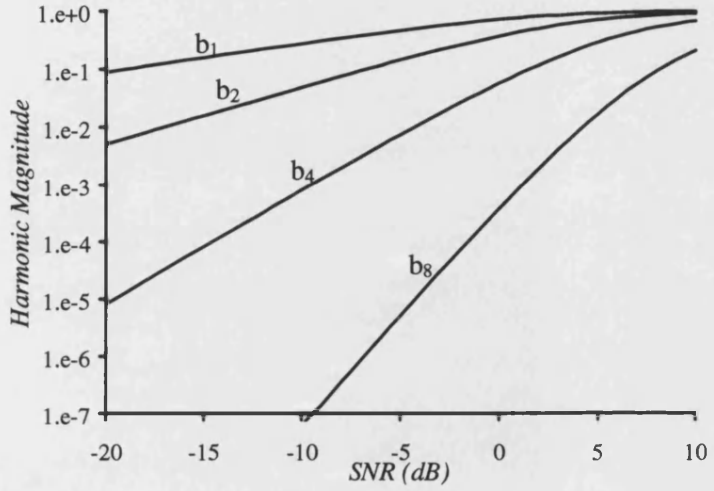


Figure 2.1: Plots of Fourier Harmonic Magnitude Against SNR

This effect can be seen visually in figure 2.2 for CW, BPSK, QPSK and 8PSK at different levels of SNR, where the more peaky distribution is at high SNR. At the lowest SNR the pdf looks sinusoidal, and can be seen to be modelled accurately by one harmonic term.

It is apparent from figure 2.1 that the higher order PSK signals have a larger spacing between adjacent harmonics at any SNR. This indicates that these will resemble more closely resemble a sinusoid at a higher SNR than that of a lower order PSK signal. This is also verified from figure 2.2, where 8PSK is seen to be sinusoidal for nearly all the SNR values examined.

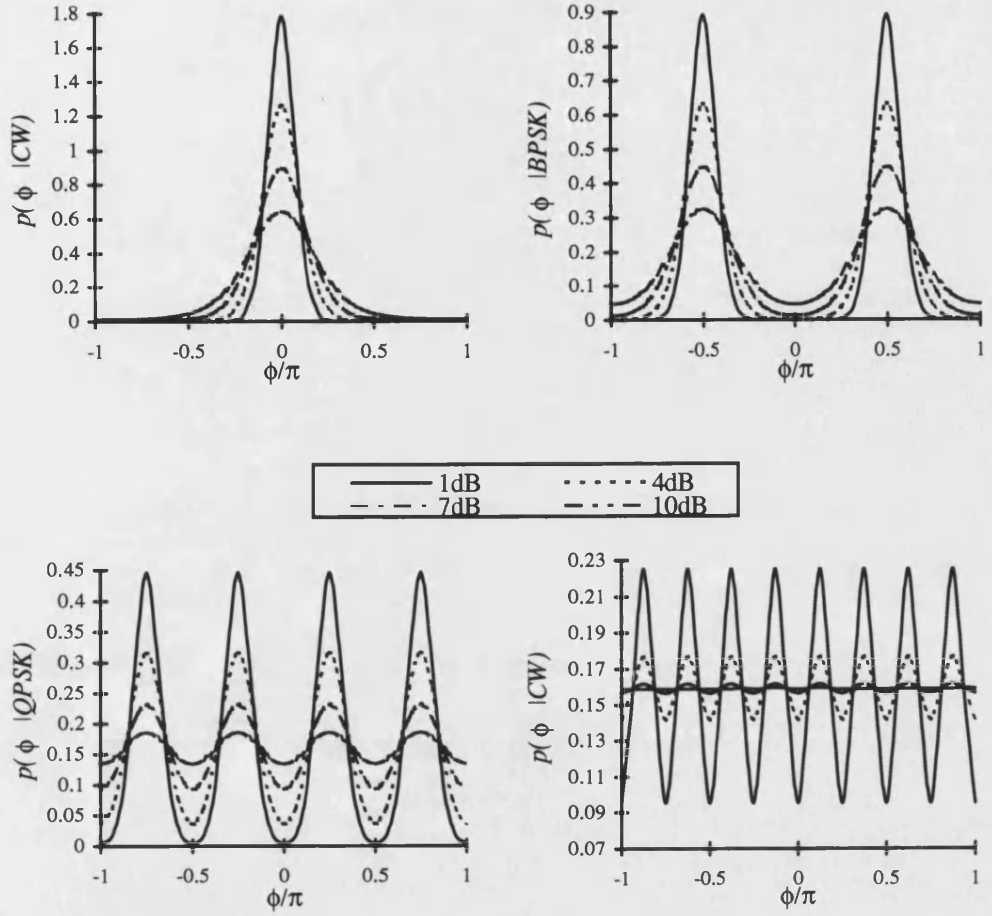


Figure 2.2 : Plots of PDF of Phase For CW, BPSK, QPSK & 8PSK at ρ=*1,4,7,10 dB SNR*

The error involved in the approximation may be quantified through the mean squared error, and is given by :

$$\varepsilon_{ms} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{1}{\pi}\sum_{n=T+1}^{\infty} b_{Mn}\cos(Mn\phi)\right]^2 d\phi \qquad (2.12)$$

This reduces to

$$\varepsilon_{ms} = \frac{1}{4\pi^2}\sum_{n=T+1}^{\infty}(b_{Mn})^2 \qquad (2.13)$$

A plot of RMS error against SNR is given in figure 2.3 for a one term approximation for CW, BPSK, QPSK and 8PSK. It can be seen from the plot that the pdf may be modelled well with a one term approximation at low SNR. It will be found in a later chapter that the approximation is excellent at the limit of detectability and is a useful tool in the analysis of PSK signals.

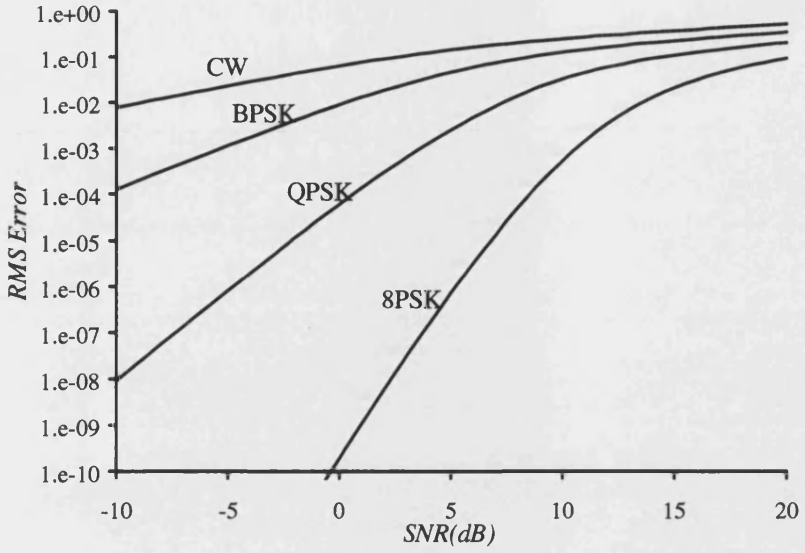


Figure 2.3 : Plots of RMS PDF Error Against SNR for A One Harmonic Term Approximation

It is therefore found from figure 2.3 that :

1. As the SNR is decreased, fewer terms are required to maintain a given level of accuracy.
2. As the PSK order is increased, fewer terms are required to maintain a given level of accuracy.
3. In a practical system e.g. SNR <30dB, the pdf may be modelled accurately with a finite number of terms.

2.3.4 High SNR development

It has been seen that the Fourier series approximation to the pdf of phase of *M*PSK signals requires more terms as the SNR is increased. At high SNR ranges, it may be appropriate to use an alternative pdf approximation. The pdf of a sinusoid in AGWN has been modelled using a Tikhonov distribution [10], where :

$$f(\phi) = \frac{\exp[2\rho\cos(\phi)]}{2\pi I_0(2\rho)} \tag{2.14}$$

This is seen to be a good approximation for a SNR>6dB, and improves with increasing SNR. The pdf may be extended to that of *M*PSK by using :

$$f_M(\phi)=\frac{1}{M}\sum_{n=0}^{M-1} f\left(\phi+\frac{\pi[2n+1]}{M}-\pi\right) \tag{2.15}$$

A final, more simple approximation may be made for SNR>14dB using a normal distribution [10], in which case the pdf of a sinusoid in AGWN is given by :

$$f(\phi)=\sqrt{\frac{\rho}{\pi}}\exp[-\rho\phi^2] \tag{2.16}$$

2.3.5 Moments of the PDF of Phase

The moments of the pdf of phase for a sinusoid in Gaussian noise are derived in Appendix 2.F and are summarised below :

$$m_n=\frac{\pi^n}{n+1}+2n!\pi^{n-2}\sum_{i=1}^{\infty} b_i\frac{(-1)^i}{i^2}\sum_{k=0}^{\frac{n}{2}-1}\frac{(-1)^k}{(n-2k-1)!(i\pi)^{2k}} \tag{2.17}$$

where m_n is the n^{th} moment. This has been extended to M level PSK [9] by :

$$m_n=\frac{\pi^n}{n+1}+2n!\pi^{n-2}\sum_{i=1}^{\infty} b_{iM}\frac{(-1)^i}{(iM)^2}\sum_{k=0}^{\frac{n}{2}-1}\frac{(-1)^k}{(n-2k-1)!(iM\pi)^{2k}} \tag{2.18}$$

2.3.6 PDF of Instantaneous Frequency

The instantaneous frequency of a signal may be defined in a number of ways, and the fundamental definition is derived from the phase $\phi(t)$:

$$f_i(t)=\frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{2.19}$$

The signal phase may be expressed in terms of the quadrature components by :

$$\phi(t)=\tan^{-1}(Q(t),I(t)) \tag{2.20}$$

and the instantaneous frequency may be deduced through a direct differentiation of (2.20) to yield :

$$f_i(t)=\frac{1}{2\pi}\frac{I(t)\frac{dQ(t)}{dt}-Q(t)\frac{dI(t)}{dt}}{I^2(t)+Q^2(t)} \tag{2.21}$$

Finally, a more unusual definition may be obtained through [19] as :

$$f_i(t) = \frac{1}{2\pi} \mathrm{Im}\left(\frac{1}{s(t)} \frac{ds(t)}{dt} \right) \tag{2.22}$$

These definitions describe continuous time relationships, and it is desirable in modulation recognition to achieve all the estimates through digital techniques.

Published work in modulation recognition has seen the use of a number of different digital instantaneous frequency operators, including estimation through Zero crossings [22][23][24][21][22], finite central difference [25], finite backward difference [26][27][28][29][30][22][32][33], the analytic approximation [19][34][35][36], autoregressive methods [37][38] and a unusual estimator, which will be termed the "Nagy" Estimator [39].

Each instantaneous frequency estimate requires at least two signal samples for analysis. This is because the estimate is approximating a derivative function (2.19). There is no fundamental upper limit on the samples in the analysis frame, but there are performance limitations which must be taken into account.

When the number of samples used in the estimate is increased, the variance of the error will decrease. A more precise relationship is found through the Cramér-Rao lower bounds [16] which determine the minimum attainable mean squared error. Much of the literature on instantaneous frequency estimation [12][13] has looked into techniques which base the estimate on a large observation window and methods are used for filtering out the noise.

The gain in performance due to an increased observation window will be achieved at the cost of bandlimiting the signal. This is a reasonable constraint when the signal does not vary rapidly with time, but with unknown signals the time varying nature can become considerable and the transitions will become smeared in the time domain.

In modulation recognition the instantaneous frequency may be composed of sharp transitions, particularly in the digital modulation schemes such as FSK. It is therefore proposed that the modulation recognition system utilises the minimum number of samples within the instantaneous frequency estimate, and that error reduction should be achieved prior to sampling in the pre-processing section through the filtering of the noise from the signal. Therefore an observation of two samples is considered to be appropriate.

Appendix 2.H outlines a comparison of the different methods of instantaneous frequency estimation detailed within the published literature on modulation recognition. It concludes that the first order backward difference approximation is a good instantaneous frequency estimator

for use in automatic modulation recognition. The estimator is found to have a good statistical performance and low computational complexity when compared with other techniques, and the statistics of this estimator are modelled for a sinusoid in AGWN within this section.

The backward difference approximation to the discrete time instantaneous frequency $f(n)$ is given by :

$$f_b(n) = \frac{1}{2\pi T_s}\left[\phi(n) - \phi(n-1)\right]_{\mathrm{mod}(2\pi)} \tag{2.23}$$

The pdf of phase difference ψ for a sinusoid in AGWN is given by the auto-convolution of the pdf of phase [3], and has been derived in [11] :

$$p_M(\psi) = \int_{-\pi}^{\pi} f(\phi) f(\psi - \phi) d\phi \tag{2.24}$$

This can be generated directly from (2.7) as :

$$p(\psi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} (b_n)^2 \cos(n\psi) \tag{2.25}$$

The range of ψ is $(-\pi,\pi]$ due to the modulo 2π arithmetic. The pdf of instantaneous frequency is derived by scaling the variable ψ by $\frac{1}{2\pi T_s}$ where T_s is the sampling frequency, and offsetting the expression by a factor of f_c.

$$p(f) = T_s + 2T_s \sum_{n=1}^{\infty} (b_n)^2 \cos\left(2\pi n\left[fT_s - f_c\right]\right) \tag{2.26}$$

The statistics of this estimator have been published in [2] by the author of this thesis. Summarising some of the important results of this paper, the mean squared error of a sinusoid of frequency f_c in AGWN is given by :

$$\mathrm{MSE} = \frac{1}{12T_s^2} + \frac{1}{\pi^2 T_s^2}\sum_{n=1}^{\infty} (b_n)^2 \frac{(-1)^2}{n^2}\cos(2\pi n f_c) + f_c\left(f_c - \frac{2}{\pi T_s}\sum_{n=1}^{\infty}(b_n)^2 \frac{(-1)^{n+1}}{n}\sin(2\pi n f_c)\right) \tag{2.27}$$

When the circular variance is analysed, the mean squared error is found to be independent of the sinusoid frequency and the mean squared error is given by :

$$MSE_{circ} = \frac{\rho - 2\ln\left[\frac{\sqrt{\pi\rho}}{2}\left(I_1\left(\frac{\rho}{2}\right) + I_0\left(\frac{\rho}{2}\right)\right)\right]}{2\pi^2 T_s^2} \tag{2.28}$$

For a detailed derivation of these results, the reader is referred to [2], a copy of which appears in Appendix III at the end of the thesis.

2.3.7 PDF of Phase Difference

The pdf of phase difference can be a useful statistic to evaluate in differential detection applications. Two cases are examined. The first is the most common, where the phase difference is applied to a modulo 2π operator. The second is where the modulo arithmetic is omitted.

The pdf of phase difference modulo 2π for a sinusoid in AGWN is given in (2.25). This may be extended to MPSK from (2.11) as :

$$p_M(\psi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}(b_{n.M})^2 \cos(nM\psi) \tag{2.29}$$

The autoconvolution of two independent phase samples when modulo 2π is not applied becomes more difficult to analyse. In this case the phase difference Ω has a range in the interval $(-2\pi,\pi]$. This has been developed in Appendix 2.G in a high and low SNR form, and is summarised as follows :

$$g(\Omega) = \frac{1}{2\pi^2}\left\{\begin{matrix}\left(\pi - \frac{\Omega}{2}\right)\left[1 + 2b_1^2\cos(\Omega) + 2b_2^2\cos(2\Omega) + 2b_3^2\cos(3\Omega)\right] + \\ k_1\sin(\Omega) + k_2\sin(2\Omega) + k_3\sin(3\Omega)\end{matrix}\right\} \quad \begin{matrix}\rho < 2\,\text{dB} \\ g(\Omega) \geq 0\end{matrix} \tag{2.30}$$

$$g(\Omega) = \frac{1}{2\pi^2\left[I_0(2\rho)\right]^2}\left\{I_0\left[4\rho\cos\left(\frac{\Omega}{2}\right)\right]\left(\pi - \frac{\Omega}{2}\right) - 2\sum_{n=1}^{8}\frac{(-1)^n}{n}I_n\left[4\rho\cos\left(\frac{\Omega}{2}\right)\right]\sin\left(n\frac{\Omega}{2}\right)\right\} \quad \begin{matrix}\rho \geq 2\,\text{dB} \\ g(\Omega) \geq 0\end{matrix} \tag{2.31}$$

where k_1, k_2 and k_3 are given in terms of b_1, b_2 and b_3 in (G.15) (G.16) and (G.17). The evaluation is applicable to positive indices, and for negative values evaluation is achieved through :

$$g(-\Omega) = g(\Omega) \tag{2.32}$$

2.4 References

[1] P.C. Sapiano, R.J. Holbeche, J.D. Martin, "Low SNR Approximation To Phase PDF For PSK Signals", IEE Electronics Letters, Vol. 30, No. 16, pp. 1279-1280, 1994.

[2] P.C. Sapiano, J.D. Martin, "Statistical Performance Of The First Order Phase Difference Digital Instantaneous Frequency Estimator", IEE Electronics Letters, to be published.

[3] A.D. Whalen, "Detection Of Signals In Noise", Academic Press, New York, Chs. 1&4, 1971.

[4] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, Washington D.C., Ch.s 6,9,10,15, 1965.

[5] D. Middleton, "An Introduction to Statistical Communication Theory", Peninsula Publishing, California, 1987, Appendix 1.

[6] K.V. Mardia, "Statistics of Directional Data", Academic Press, London, 1972, Ch. 3.

[7] N.M. Blachman, "Gaussian Noise-Part II: Distribution Of Phase Change Of Narrow-Band Noise Plus Sinusoid", IEEE Trans. Information Theory, Vol. 34, No. 6, pp. 1401-1405, November 1988.

[8] I. S. Gradshteyn and I.M. Ryzhik, "Table Of Integrals And Series Products", Academic Press, London, 1980.

[9] Y. Yang and S.S. Soliman, "Statistical Moments Based Classifier For MPSK Signals", Proc. IEEE GLOBECOM, pp. 2.7.1-2.7.5, 1991.

[10] H. Leib and S. Pasupathy, "The Phase Of A Vector Perturbed By Gaussian Noise And Differentially Coherent Detectors", IEEE Trans. Information Theory, Vol. 34, No.6, pp. 1491-1501, November 1988.

[11] R.F. Pawula, S.O. Rice, J.H. Roberts, "Distribution of the Phase Angle Between Two Vectors Perturbed by Gaussian Noise", IEEE Trans. Comms., Vol. 30, No. 8, Aug. 1982, pp. 1828-1841.

[12] B. Boashash, "Estimating and Interpreting the Instantaneous Frequency of a Signal - Part 2 : Algorithms and Applications", Proceedings of the IEEE, Vol. 80, No. 4, pp.540-568, 1992.

[13] M. Sun, R.J. Sclabassi, "Discrete-Time Instantaneous Frequency and its Computation", IEEE Trans. Signal Processing, Vol. 41, No. 5, pp. 1867-1879.

[14] B.C. Lovell, R.C. Williamson, "The Statistical Performance of some Instantaneous Frequency Estimators", IEEE Trans. Signal Processing, Vol. 40, No. 7, July 1992, pp. 1708-1723.

[15] N.M. Blachman, "Gaussian Noise-Part II: Distribution Of Phase Change Of Narrow-Band Noise Plus Sinusoid", IEEE Trans. Information Theory, Vol. 34, No. 6, pp. 1401-1405, November 1988.

[16] D.C. Rife, R.R. Boorstyn, "Single-Tone Parameter Estimation from Discrete-Time Observations", IEEE trans. Information Theory, Vol. IT-20, No.5, Sept. 1974, pp. 591-598

[17] K.V. Mardia, "Statistics of Directional Data", Academic Press, London, 1972, Ch. 2&3

[18] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, Washington D.C., Ch. 9, 1965.

[19] P.A.J. Nagy, "Analysis of a Method for Classification of Analogue Modulated Radio Signals", EUSIPCO '94.

[20] T.G. Callaghan, J.L. Perry, J.K. Tjho, "Sampling and Algorithms aid Modulation Recognition", Microwaves & R.F., September 1985, pp. 117-121.

[21] J. Portillo Garcia, J.P. Sancho-Marco, L. Vergara Dominguez, J.M. Páez Borrallo, Ruiz-Mezcua, "A Microcomputer-Based General Architecture for Radiocommunication Signal Classification and Digital Demodulation", Proc. EUSIPCO '90, vol. 3, pp. 1919-1922.

[22] L. Vergara Dominguez, J.M. Páez Borrallo, J. Portillo Garcia, "A General Approach to the Automatic Classification of Radiocommunication signals", Signal Processing, Vol. 22, 1991, pp. 239-250.

[23] S.B. McMillan, B.P. Flanagan, T.K. Doong, "Determination of the Modulation Type of Communication Signals", IEEE ICASSP'90, pp. 1683-1686.

[24] S.Z. Hsue, S.S. Soliman, "Automatic Modulation Classification Using Zero Crossing", IEE Proceedings, Vo. 137, Pt. F, No. 6, Dec. 1990, pp. 459-464.

[25] F.F. Liedtke, "Computer Simulation of an Automatic Classification Procedure for Digitally Modulated Communication Signals with Unknown Parameters", Signal Processing, Vol. 6, No. 4, August 1984, pp. 311-323.

[26] C. Louis, P. Sehier, "Automatic Modulation Recognition with a Hierarchical Neural Network", MILCOM '94, pp. 713-717.

[27] E.R. Adams, P.C.J. Hill, C.N. Kempson, "A Statistical Approach to Modulation Recognition for Communications Signal Monitoring and Surveillance", IERE Sept. 1988, pp. 31-37.

[28] P.C.J. Hill, G.R. Orzeszko, "Performance Comparison of Neural Network and Statistical Discriminant Processing Techniques for Automatic Modulation Recognition", Proc. SPIE, Vol. 1469, Applications of Artificial Neural Networks, 1991, pp. 329-340.

[29] J.E. Hipp, "Modulation Classification Based on Statistical Moments", IEEE MILCOM '86, pp. 20.2.1-20.2.6.

[30] P.M. Fabrizi, L.B. Lopes, G.B. Lockhart, "Receiver Recognition of Analogue Modulation Type", IERE Radio Receivers and Associated Systems '86, pp. 135-140.

[31] L. Vergara Dominguez, J.M. Páez Borrallo., J. Portillo Garcia, "A General Approach to the Automatic Classification of Radiocommunication signals", Signal Processing, Vol. 22, 1991, pp. 239-250.

[32] J.E. Whelchel, D.L. McNeill, R.D. Hughes, M.M. Loos, "Signal Understanding : an Artificial Intelligence Approach to Modulation Classification", Proc. IEEE International Workshop on Tools for A.I. 1989, pp. 231-236.

[33] F. Jondral, "Automatic Classification of High Frequency Signals", Signal Processing, Vol. 9, No. 3, October 1985, pp. 177-190.

[34] R.J. Mammone, R.J. Rothaker, C.I. Podilchuk, S. Davidovici, D.L. Schilling, "Estimation of Carrier Frequency, Modulation Type and Bit Rate of a Modulation Signal", IEEE ICC '87, pp. 28.4.1- 28.4.7.

[35] K.B. Parcell, J. Aisbett, "Automatic Modulation", Signal processing, ISSPA' 87, vol. 2, pp. 871-874.

[36] G.A. Einicke, "A Flexible Architecture for Real-time Modulation Recognition", ASSPA '89, pp. 133-137.

[37] K. Assaleh, K. Farrell, Mammone, R.J., "A New Method of Modulation Classification for Digitally Modulated Signals", MILCOM '92, pp. 30.5.1-30.5.5.

[38] K.R. Farrell, R.J. Mammone, "Modulation Classification Using a Neural Tree Network", MILCOM '93, pp.1028-1032.

[39] P.A.J. Nagy, "A Modulation Classifier for Multi Channel Systems and Multi Transmitter Situations", IEEE MILCOM '94, Vol. 3, pp. 816-820.

[40] K.P. Zimmerman, "On Frequency-Domain and Time-Domain Phase Unwrapping", Proceedings of the IEEE, Vol. 75, No. 4, April 1987, pp. 519-520.

[41] J.M. Tribolet, "A New Phase Unwrapping Algorithm", IEEE Trans. ASSP, Vol. 25, No. 2, April 1977, pp. 170-177.

[42] W. Martin, P. Flandrin, "Wigner-Ville Spectral Analysis of Nonsationary Processes", IEEE Trans. ASSP, Vol. 33, No. 6, December 1985, pp. 1461-1470.

[43] H.L. Van Trees, "Detection, Estimation and Modulation Theory", Part I, New York, Wiley, 1968.

[44] K.V. Mardia, "Statistics of Directional Data", Academic Press, London, 1972, Ch. 2&3.

2.5 Appendix 2.A

Proof of the Fourier Series Expansion of a Sinusoid in AGWN

The following proof develops the Fourier series representation for the pdf of phase given by :

$$f(\phi)=\frac{1}{2\pi}e^{-\rho}+\frac{1}{2}\sqrt{\frac{\rho}{\pi}}e^{-\rho}\cos(\phi)e^{\rho\cos^2(\phi)}\left[1+\mathrm{erf}\left(\sqrt{\rho}\cos(\phi)\right)\right] \tag{A.1}$$

The Fourier series form is given in (2.7) as

$$f(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{m=1}^{\infty}b_m\cos(m\phi) \tag{A.2}$$

where

$$b_m=\int_{-\pi}^{\pi}f(\phi)\cos(m\phi)d\phi \tag{A.3}$$

In order to evaluate this, the following integral must be performed :

$$\int_{-\pi}^{\pi}\cos\phi\cos(m\phi)\exp\left[\rho\cos^2\phi\right]\left(1+erf\left[\sqrt{\rho}\cos\phi\right]\right)d\phi \tag{A.4}$$

This integral is split into two parts :

$$I=\int_{-\pi}^{\pi}\cos\phi\cos(m\phi)\exp\left[\rho\cos^2\phi\right]erf\left[\sqrt{\rho}\cos\phi\right]d\phi \tag{A.5}$$

$$J=\int_{-\pi}^{\pi}\cos\phi\cos(m\phi)\exp\left[\rho\cos^2\phi\right]d\phi \tag{A.6}$$

The error function may be expressed in series form by [4] :

$$erf(z)=\frac{1}{\sqrt{\pi}}e^{-z^2}\sum_{n=0}^{\infty}\frac{2^{2n+2}z^{2n+1}(n+1)!}{(2n+2)!} \tag{A.7}$$

Therefore :

$$\cos\phi\cos(m\phi)\exp\left[\rho\cos^2\phi\right]erf\left[\sqrt{\rho}\cos\phi\right]=\frac{1}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{2^{2n+2}\rho^{n+\frac{1}{2}}\cos m\phi(\cos\phi)^{2n+2}(n+1)!}{(2n+2)!} \tag{A.8}$$

Note the following expansion [8]

$$\cos(m\phi)(\cos\phi)^{2n+2} = \frac{1}{2^{2n+2}}\left\{\sum_{k=0}^{n} 2\binom{2n+2}{k}\cos(m\phi)\cos[2(n+1-k)\phi] + \binom{2n+2}{n}\cos(m\phi)\right\} \quad \text{(A.9)}$$

Therefore through orthogonality

$$\int_{-\pi}^{\pi}\cos(m\phi)(\cos\phi)^{2n+2}\,d\phi = \begin{cases} \dfrac{1}{2^{2n+1}}\dfrac{\pi(2n+2)!}{(n+1-\frac{m}{2})!(n+1+\frac{m}{2})!} & \text{Even m} \quad n \geq \frac{m-2}{2} \\ 0 & \text{Odd m} \end{cases} \quad \text{(A.10)}$$

With some manipulation :

$$I = \begin{cases} 2\sqrt{\rho\pi}\,\rho^{\frac{m}{2}-1}\sum_{k=0}^{\infty}\dfrac{\Gamma(k+\frac{m}{2}+1)}{k!\Gamma(k+m+1)}\rho^{k} & \text{Even m} \\ 0 & \text{Odd m} \end{cases} \quad \text{(A.11)}$$

Now looking at the second part of the integral :

$$J = \int_{-\pi}^{\pi}\cos\phi\cos(m\phi)\exp[\rho\cos^2\phi]d\phi = \sum_{n=0}^{\infty}\frac{\rho^n}{n!}\int_{-\pi}^{\pi}\cos(m\phi)(\cos\phi)^{2n+1}d\phi \quad \text{(A.12)}$$

Using

$$\cos(m\phi)(\cos\phi)^{2n+1} = \frac{1}{2^{2n}}\sum_{k=0}^{n}\binom{2n+1}{k}\cos(m\phi)\cos[(2n+1-2k)\phi] \quad \text{(A.13)}$$

It can be seen that :

$$\int_{-\pi}^{\pi}\cos(m\phi)(\cos\phi)^{2n+2} = \begin{cases} \dfrac{1}{2^{2n}}\dfrac{\pi(2n+1)!}{(n+\frac{1}{2}-\frac{m}{2})!(n+\frac{1}{2}+\frac{m}{2})!} & \text{Odd m} \quad n \geq \frac{m-1}{2} \\ 0 & \text{Even m} \end{cases} \quad \text{(A.14)}$$

Using

$$\frac{(2n+1)!}{(n)!} = \frac{2^{2n+1}\Gamma(n+1+\frac{1}{2})}{\sqrt{\pi}} \quad \text{(A.15)}$$

It is found that :

$$J = \begin{cases} 2\sqrt{\pi\rho}\,\rho^{\frac{m}{2}-1}\sum_{k=0}^{\infty}\dfrac{\Gamma(\kappa+\frac{m}{2}+1)}{\kappa!\Gamma(\kappa+m+1)}\rho^{k} & \text{Odd } m \\ 0 & \text{Even } m \end{cases} \quad \text{(A.16)}$$

which is the same result as (A.11), but for an odd *m*. Combining (A.11) and (A.16) it is seen that :

$$b_m = e^{-\rho}\rho^{\frac{m}{2}}\sum_{k=0}^{\infty}\frac{\Gamma\left(k+\frac{m}{2}+1\right)}{k!\Gamma(k+m+1)}\rho^k \tag{A.17}$$

Using the definition for the Confluent Hypergeometric function :

$${}_1F_1(a;b;z) = \frac{\Gamma(b)}{\Gamma(a)}\sum_{n=0}^{\infty}\frac{\Gamma(n+a)z^n}{\Gamma(n+b)n!} \tag{A.18}$$

The function (A.17) may be simplified to

$$b_m = e^{-\rho}\rho^{\frac{m}{2}}\frac{\Gamma\left(\frac{m}{2}+1\right)}{\Gamma(m+1)}{}_1F_1\left(\tfrac{m}{2}+1;m+1;\rho\right) \tag{A.19}$$

Using the Kummer transformation [4] this simplifies to

$$b_m = \rho^{\frac{m}{2}}\frac{\Gamma\left(\frac{m}{2}+1\right)}{\Gamma(m+1)}{}_1F_1\left(\tfrac{m}{2};m+1;\rho\right) \tag{A.20}$$

The relationship in terms of Bessel functions may be found by observing that :

$$J = \frac{e^{\frac{\rho}{2}}}{2}\int_{-\pi}^{\pi}\exp\left[\tfrac{\rho}{2}\cos(2\phi)\right]\left(\cos(m+1)\phi+\cos(m-1)\phi\right)d\phi \tag{A.21}$$

and using the identity

$$I_n(z) = \frac{1}{\pi}\int_0^{\pi}e^{z\cos\theta}\cos(n\theta)d\theta \tag{A.22}$$

Comparing (A.11) and (A.16) it is found that

$$b_m = \frac{\sqrt{\rho\pi}\,e^{-\frac{\rho}{2}}}{2}\left[I_{\frac{m+1}{2}}\left(\tfrac{\rho}{2}\right)+I_{\frac{m-1}{2}}\left(\tfrac{\rho}{2}\right)\right] \tag{A.23}$$

2.6 Appendix 2.B

Properties of the Even Fourier Harmonics

Bessel functions of integer order plus a half are related to the Modified Spherical Bessel function [4]. The following recurrence property is useful for evaluation :

$$I_{v+1}(x) = -\frac{2v}{x} I_v(x) + I_{v-1}(x) \tag{B.1}$$

Using the expression [4] :

$$I_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sinh(x) \tag{B.2}$$

$$I_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \cosh(x) \tag{B.3}$$

The following properties may be deduced :

$$\frac{\sqrt{\rho\pi}\, e^{-\frac{\rho}{2}}}{2} I_{\frac{1}{2}}\left(\frac{\rho}{2}\right) = \frac{1-\exp(-\rho)}{2} \tag{B.4}$$

This property is used later in Appendix 2.C. Some final simple calculations yield the following :

$$b_2 = \frac{\rho + \exp(-\rho) - 1}{\rho} \tag{B.5}$$

$$b_4 = \frac{\rho^2 - 4\rho + 6 - 2[\rho + 3]\exp(-\rho)}{\rho^2} \tag{B.6}$$

2.7 Appendix 2.C

Recursive Algorithm for Fourier Harmonic Generation

The following work discusses an efficient routine for calculating all of the Fourier coefficients. The evaluation of Bessel functions is usually performed using the recurrence relationship of (B.1), which is valid for integer order and integer plus a half order Bessel functions.

This process is upwardly unstable in terms of error and therefore, reverse recursion is generally used. With reverse recursion, arbitrary values are used to specify a high order Bessel function, and the algorithm is iterated downwards to a low order function.

The error reduces through each iteration, until the smallest order Bessel function (order >0) is reached. This smallest order Bessel function is determined using numerical approximations and a normalisation constant is determined by looking at the ratio of the approximation to the iterate. All the iterative values are then normalised by this constant.

The algorithm for Fourier coefficient generation is split into the Bessel functions of integer order and integer plus a half order. The algorithm is detailed below :

Algorithm for Fourier Coefficient Generation

1. Formulae for iteration : $\tilde{I}o_{n-1} = \tilde{I}o_{n+1} + \frac{4n}{\rho}\tilde{I}o_n \qquad \tilde{I}e_{(n+\frac{1}{2}-1)} = \tilde{I}e_{(n+\frac{1}{2}+1)} + \frac{4n+2}{\rho}\tilde{I}e_{(n+\frac{1}{2})}$
2. Set a maximum Bessel order *MAX* e.g. 80
3. Set $\tilde{I}e_{MAX} = \tilde{I}o_{MAX+\frac{1}{2}} = 0$ and $\tilde{I}e_{MAX-1} = \tilde{I}o_{MAX-\frac{1}{2}} = 1$
4. Perform the recursions of line 1 for n=*MAX*-1...0
5. Determine $ro = \frac{\sqrt{\rho\pi}\exp\left(-\frac{\rho}{2}\right)I_0\left(\frac{\rho}{2}\right)}{2\tilde{I}o_0}$ through a numerical approximation [4], and $re = \frac{1-\exp(-\rho)}{2\tilde{I}e_0}$ from (B.4)
6. Perform $\ell o_n = ro\,\tilde{I}o_n$ and $\ell e_{n+\frac{1}{2}} = re\,\tilde{I}e_{n+\frac{1}{2}}$ for n=*MAX*-1...0
7. The final result is given by $b_{2n+1} = \ell o_n + \ell o_{n+1} \qquad b_{2n+2} = \ell e_{n+\frac{1}{2}} + \ell e_{n+\frac{1}{2}+1}$ for n=0...*MAX*-1

Therefore a whole set of Bessel coefficients are generated through a simple iterative algorithm.

2.8 Appendix 2.D

Iterative Algorithm for Generating Cosine and Sine of Multiple Angles

Using the following trigonometric identities :

$$\cos[(n+1)\phi] = \cos(n\phi)\cos(\phi) - \sin(n\phi)\sin(\phi) \tag{D.1}$$

and

$$\sin[(n+1)\phi] = \sin(n\phi)\cos(\phi) - \cos(n\phi)\sin(\phi) \tag{D.2}$$

and using the notation

$$c_0 = \cos(\phi) \tag{D.3}$$

and

$$s_0 = \sin(\phi) \tag{D.4}$$

a recurrence relationship may be deduced for generating the cosine and sine of a series of multiple angles. Using the recurrence relationship below :

$$c_{n+1} = c_n c_0 - s_n s_0 \tag{D.5}$$

$$s_{n+1} = s_n c_0 + c_n s_0 \tag{D.6}$$

and for the range $n = \{0,1,2...L-1\}$ where L is the maximum number coefficients, the set of cosine terms is given by :

$$\cos(n\phi) = c_{n-1} \tag{D.7}$$

2.9 Appendix 2.E

Fourier Series Expansion of the PDF of Phase for *M*PSK

The pdf of phase of a M level PSK signal in the presence of Gaussian noise, may be developed from the pdf of phase of a sinusoid, as follows :

$$f_M(\phi)=\frac{1}{M}\sum_{n=0}^{M-1} f(\phi-\theta_n) \tag{E.1}$$

where

$$\theta_n=\frac{\pi[2n+1]}{M}-\pi \tag{E.2}$$

This is developed from (2.7) as :

$$f_m(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{i=1}^{\infty} b_i \frac{1}{M}\sum_{n=0}^{M-1}\cos\left(i[\phi-\theta_n]\right) \tag{E.3}$$

The internal part of (E.3) may be expressed as follows :

$$\sum_{n=0}^{M-1}\cos\left(i[\phi-\theta_n]\right)=\mathrm{Re}\left\{e^{j(i\phi)}\sum_{n=0}^{M-1}\left(e^{-j\theta_n}\right)^i\right\} \tag{E.4}$$

and by analysis of the phase vectors of PSK and the effect of phase rotation due to the power term, it can be seen that :

$$\sum_{n=0}^{M-1}\cos\left(i[\phi-\theta_n]\right)=\begin{cases}\mathrm{Re}\left\{e^{j(i\phi)}M(-1)^{\frac{i}{M}}\right\} & i=Mm\\ 0 & i\neq Mm\end{cases} \tag{E.5}$$

where m is an integer. Therefore the pdf of phase for *M*PSK may be written as :

$$f_m(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^n b_n\cos(nM\phi) \tag{E.6}$$

2.10 Appendix 2.F

Statistical Moments of the Phase of *M*PSK

The n^{th} moment of the pdf of phase of a signal is given by :

$$m_n = \int_{-\pi}^{\pi} \phi^n f_M(\phi) d\phi \tag{F.1}$$

Using the identity from [8, eqn. 3.76.10] it is found that :

$$\int_{-\pi}^{\pi} \phi^n \cos(m\phi) d\phi = 2\frac{(-1)^m}{m^{n+1}} \sum_{k=0}^{\frac{n}{2}-1} \frac{(-1)^k n!(m\pi)^{n-2k-1}}{(n-2k-1)!} \tag{F.2}$$

Using this in (2.7), the phase moments of a sinusoid in AGWN are given by :

$$m_n = \frac{\pi^n}{n+1} + 2n!\pi^{n-2} \sum_{i=1}^{\infty} b_i \frac{(-1)^i}{i^2} \sum_{k=0}^{\frac{n}{2}-1} \frac{(-1)^k}{(n-2k-1)!(i\pi)^{2k}} \tag{F.3}$$

and for *M* level PSK this is given through (2.11) as :

$$m_n = \frac{\pi^n}{n+1} + 2n!\pi^{n-2} \sum_{i=1}^{\infty} b_{iM} \frac{(-1)^i}{(iM)^2} \sum_{k=0}^{\frac{n}{2}-1} \frac{(-1)^k}{(n-2k-1)!(iM\pi)^{2k}} \tag{F.4}$$

2.11 Appendix 2.G

PDF of Phase Difference Without 2π Wrapping

The auto-convolution when modulo 2π is not used, takes the form :

$$g(\Omega) = \int_{-2\pi}^{2\pi} f(y) f(\Omega - y) dy \tag{G.1}$$

and for $\Omega > 0$ this may be expressed as

$$g(\Omega) = \int_{\Omega-\pi}^{\pi} f(y) f(\Omega - y) dy \tag{G.2}$$

and $g(-\Omega) = g(\Omega)$

2.11.1.1 Fourier Series Form

Using the Fourier series representation for the pdf of phase (2.7) with $b_0 = 0.5$, it can be shown that

$$g(\Omega) = \frac{1}{\pi^2} \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} b_i b_n \int_{\Omega-\pi}^{\pi} \cos\big(i(\Omega - y)\big) \cos(ny) dy \tag{G.3}$$

For the case where $i \neq n$ it can be shown that the integral term is given by

$$(-1)^{i+n+1} \frac{n \sin(n\Omega) - i \sin(i\Omega)}{n^2 - i^2} \tag{G.4}$$

In the case where $i = n \neq 0$ the integral term is given by

$$\left(\pi - \tfrac{\Omega}{2}\right) \cos(n\Omega) - \frac{\sin(n\Omega)}{2n} \tag{G.5}$$

Finally for $i = n = 0$ the integral term is given by

$$\left(\pi - \tfrac{\Omega}{2}\right) \tag{G.6}$$

Combining all three cases

$$g(\Omega)=\left(\pi-\tfrac{\Omega}{2}\right)\frac{b_0^2}{\pi^2}+\frac{1}{\pi^2}\sum_{n=1}^{\infty}b_n^2\left(\pi-\tfrac{\Omega}{2}\right)\cos(n\Omega)-\frac{1}{\pi^2}\sum_{n=1}^{\infty}b_n^2\frac{\sin(n\Omega)}{2n}+\ldots$$
$$\frac{1}{\pi^2}\sum_{n=0}^{\infty}\sum_{\substack{i=0\\ i\neq n}}^{\infty}(-1)^{i+n+1}b_ib_n\frac{n\sin(n\Omega)-i\sin(i\Omega)}{n^2-i^2} \qquad \text{(G.7)}$$

Using the following identity :

$$\sum_{n=0}^{\infty}\sum_{i=0}^{\infty}(-1)^{i+n+1}b_ib_n\frac{n\sin(n\Omega)-i\sin(i\Omega)}{n^2-i^2}=2\sum_{n=0}^{\infty}\left\{n(-1)^nb_n\sin(n\Omega)\sum_{i=0}^{\infty}\frac{b_i(-1)^i}{i^2-n^2}\right\} \qquad \text{(G.8)}$$

(G.7) can be simplified to :

$$g(\Omega)=\frac{1}{2\pi^2}\left\{\left(\pi-\tfrac{\Omega}{2}\right)\left[1+2\sum_{n=1}^{\infty}b_n^2\cos(n\Omega)\right]+\sum_{n=1}^{\infty}nb_n\sin(n\Omega)\left[4(-1)^n\left[\sum_{\substack{i=0\\ i\neq n}}^{\infty}\frac{(-1)^ib_i}{i^2-n^2}\right]-\frac{b_n}{n^2}\right]\right\} \qquad \text{(G.9)}$$

2.11.1.2 Tikhonov Approximation

Using the Tikhonov approximation (2.14) a high SNR form of the pdf may be written as :

$$g(\Omega)=\frac{1}{4\pi^2\left[I_0(2\rho)\right]^2}\int_{\Omega-\pi}^{\pi}\exp\left[4\rho\cos\left(\tfrac{\Omega}{2}\right)\cos\left(y-\tfrac{\Omega}{2}\right)\right]dy \qquad \text{(G.10)}$$

By a change of variable (G.10) may be expressed as :

$$g(\Omega)=\frac{1}{4\pi^2\left[I_0(2\rho)\right]^2}\int_{-\left(\pi-\frac{\Omega}{2}\right)}^{\pi-\frac{\Omega}{2}}\exp\left[4\rho\cos\left(\tfrac{\Omega}{2}\right)\cos(x)\right]dx \qquad \text{(G.11)}$$

Using the expansion in[4] :

$$g(\Omega)=\frac{2}{4\pi^2\left[I_0(2\rho)\right]^2}\int_{0}^{\pi-\frac{\Omega}{2}}\left\{I_0\left[4\rho\cos\left(\tfrac{\Omega}{2}\right)\right]+2\sum_{n=1}^{\infty}I_n\left[4\rho\cos\left(\tfrac{\Omega}{2}\right)\right]\cos(nx)\right\}dx \qquad \text{(G.12)}$$

Finally through integration and simplification :

$$g(\Omega)=\frac{1}{2\pi^2\left[I_0(2\rho)\right]^2}\left\{I_0\left[4\rho\cos\left(\tfrac{\Omega}{2}\right)\right]\left(\pi-\tfrac{\Omega}{2}\right)-2\sum_{n=1}^{\infty}\frac{(-1)^n}{n}I_n\left[4\rho\cos\left(\tfrac{\Omega}{2}\right)\right]\sin\left(n\tfrac{\Omega}{2}\right)\right\} \qquad \text{(G.13)}$$

2.11.1.3 Combined Approach

The expression in (G.9) consists of a two dimensional infinite summation, and truncation is essential. The truncation will cause the model to be accurate at a low SNR. The pdf of phase difference may then be modelled for a full range of SNR using the two approximations, where the first three terms of (G.9) are used for a low SNR approximation and around eight terms of (G.13) for a high SNR approximation.

The results of the low SNR (ρ<2dB) approximation are

$$g(\Omega)=\frac{1}{2\pi^2}\left\{\begin{matrix}\left(\pi-\frac{\Omega}{2}\right)\left[1+2b_1^2\cos(\Omega)+2b_2^2\cos(2\Omega)+2b_3^2\cos(3\Omega)\right]+\\ k_1\sin(\Omega)+k_2\sin(2\Omega)+k_3\sin(3\Omega)\end{matrix}\right\} \tag{G.14}$$

$$k_1=\frac{b_1}{6}\left(12-6b_1-8b_2+3b_3\right) \tag{G.15}$$

$$k_2=\frac{b_2}{30}\left(-30+80b_1-15b_2-48b_3\right) \tag{G.16}$$

$$k_3=\frac{b_3}{30}\left(20-45b_1+72b_2-10b_3\right) \tag{G.17}$$

and the high SNR approximation (ρ≥2dB)

$$g(\Omega)=\frac{1}{2\pi^2\left[I_0(2\rho)\right]^2}\left\{I_0\left[4\rho\cos\left(\frac{\Omega}{2}\right)\right]\left(\pi-\frac{\Omega}{2}\right)-2\sum_{n=1}^{8}\frac{(-1)^n}{n}I_n\left[4\rho\cos\left(\frac{\Omega}{2}\right)\right]\sin\left(n\frac{\Omega}{2}\right)\right\} \tag{G.18}$$

The switch over threshold of 2dB has been chosen experimentally, and results in low peak error. This is shown graphically in figure 2.4, where the pdf and its approximation are shown in figure 2.4a and the error with the true pdf in figure 2.4b, where ρ=2dB.

It is seen that the model is good at this point, which is the result of maximum error.

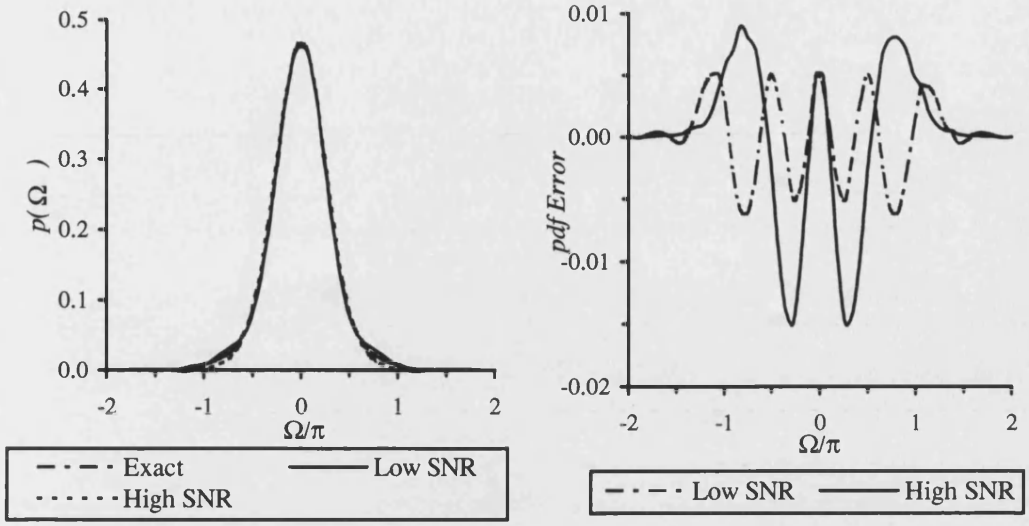


Figure 2.4 : Plots of (a) PDF of Phase Difference and (b) PDF Error

2.12 Appendix 2.H

Comparison of Instantaneous Frequency Estimation Techniques

In automatic modulation recognition the instantaneous frequency has been recognised to be a key parameter used to characterise frequency modulated signals. The instantaneous frequency of a signal may be derived in a number of ways, but all methods are merely an approximation to the true value.

This appendix outlines some of the methods which have been used to estimate the instantaneous frequency in the modulation recognition literature, and discusses the performance and complexity of each of them in turn. A comparison of the methods is made and the first order backward difference approximation is chosen as a suitable method for instantaneous frequency estimation in a modulation recognition environment.

Some results from the analysis in this section have been published in [2].

2.12.1 Instantaneous Frequency Estimators

2.12.1.1 First order difference approximation, mod 2π

The simplest approximation to the definition in (2.19) is achieved by approximating the differentiation of phase through a two point differential approximation :

$$f_c(n) = \frac{1}{4\pi T_s}\left[\phi(n+1) - \phi(n-1)\right]_{\text{mod}(2\pi)} \tag{H.1}$$

The modulo 2π is introduced to eliminate the effects due to phase wraps. This expression is often referred to as the central finite difference [12]. A more accurate approximation is given by the backward finite difference, expressed as :

$$f_b(n) = \frac{1}{2\pi T_s}\left[\phi(n) - \phi(n-1)\right]_{\text{mod}(2\pi)} \tag{H.2}$$

The backward difference operator is often not used because of the group delay which is introduced through the half sample displacement [12]. However, this reasoning is not always well justified, as there are often over-riding benefits associated with the use of the backward finite difference operator. The first benefit is that the error in terms of a derivative function is smaller than the central difference. The second benefit is that the maximum instantaneous frequency which can be measured, is at the Nyquist rate, whereas the central finite difference operator measurements may only be measured up to half the Nyquist rate.

These difference operators are often associated with the effects of phase wrapping and frequency spikes. But this only occurs when the modulo 2π is not used. The reason for using modulo 2π arithmetic is seen when the equation H.2 is written in the following form :

$$f_b(n) = \frac{1}{2\pi T_s} \arg[s(n)\bar{s}(n-1)] \tag{H.3}$$

This is essentially the angle between adjacent analytic vectors, which has been scaled in order to map it to the instantaneous frequency. Equation (H.3) provides a computationally efficient method for instantaneous frequency estimation.

2.12.1.2 Zero crossing estimate

The zero crossing method is one of the simplest forms of instantaneous frequency estimation. The technique counts the number of zeros (k) which occur during a time interval T_z and estimates the instantaneous frequency by :

$$f_z(n) = \frac{k}{2T_z} \tag{H.4}$$

It has been shown [12] that a signal with k zeros within the analysis time frame is equivalent to the average of k adjacent backward finite difference estimators, as in section 2.12.1.1.

This technique is subject to quantisation noise, which results from the fact that the number of cycles within the observation period is rounded down to the nearest integer. This is particularly prevalent at lower frequencies where there are a low number of cycles occurring within the observation period.

The method is not useful with digitally sampled signals, as the observation window length is required to be greater than two. The technique is only of any real use when estimation is performed on the analogue signal, and the result is then passed to the processor.

2.12.1.3 Analytic approximation

This technique has been derived from direct differentiation of the phase expression in terms of in-phase and quadrature components, and is then expressed in discrete form from (2.21) as :

$$f_a(n) = \frac{1}{2\pi T_s} \frac{I(n)Q'(n) - Q(n)I'(n)}{I^2(n) + Q^2(n)} \tag{H.5}$$

where the $'$ operator represents the derivative with respect to time. It is noticed that this result does not require the implicit evaluation of the signal phase.

The method has often been chosen because it does not incur phase wrapping, which was noticed in [40] where the result was derived using an alternative route through the application of a classic phase unwrapping algorithm [41] to the time domain. The invariance to phase wrapping is not a good reason to choose this method, as it was shown in section 2.12.1.1 that phase wrapping is not a real problem with the backward difference operator when modulo 2π arithmetic is used on the phase difference.

The denominator of (H.5) is seen to be the instantaneous envelope of the signal and the numerator requires a mixture of in-phase, quadrature and associated derivatives. The problem here is in the evaluation of the derivative terms. This is not quite as straight forward as the evaluation in section 2.12.1.1, as a group delay in the differentiator will mis-align other terms in the calculation.

The differential approximation has been achieved using a five point Taylor series approximation in [36] :

$$g_5'(n) \approx \frac{1}{12T_s}\left[g(n-2) - 8g(n-1) + 8g(n+1) - g(n+2)\right] \tag{H.6}$$

and in a DFT form in [19] :

$$g_f'(n) \approx \mathrm{IDFT}\left(jG(k)\frac{2\pi k}{N}\right) \tag{H.7}$$

where N is the number of samples, and $G(k)$ is the discrete Fourier transform of the signal $g(n)$. This then gives a frequency domain approximation to the differentiation process.

The most common method for obtaining the differentiation approximation is through the central difference operator :

$$g_1'(n) \approx \frac{1}{2T_s}\left[g(n+1) - g(n-1)\right] \tag{H.8}$$

It has been found that the estimation of the derivative function using (H.6) and (H.8), will result in poor results when the instantaneous frequency is above a certain value. This is due to the fact that the differentiators have associated frequency responses, which are good approximations at low frequency but poor at high frequency. The associated frequency responses may be determined by taking the differentiator approximations into the frequency domain, to leave :

$$G_1(\omega) = j\sin(\omega T_s) \tag{H.9}$$

$$G_5(\omega) = \tfrac{j}{3}\sin(\omega T_s)\left[4 - \cos(\omega T_s)\right] \tag{H.10}$$

The responses of (H.9) and (H.10) are plotted in figure 2.5.

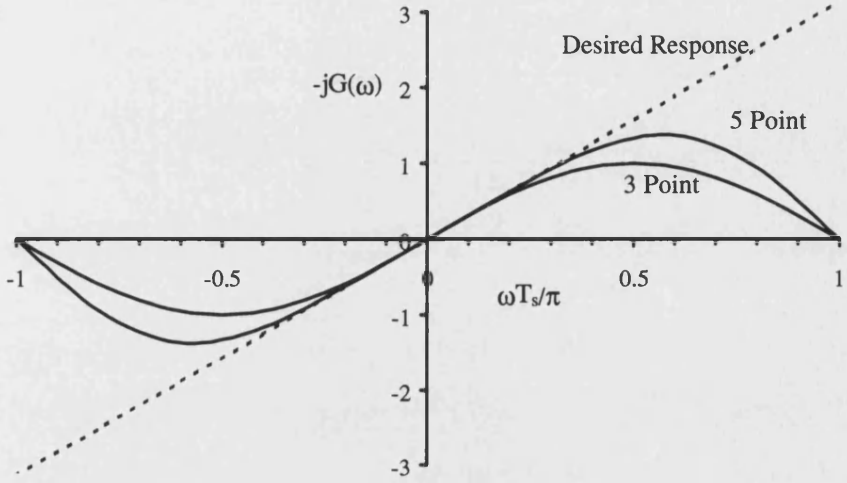


Figure 2.5 : Frequency Responses of a Three Point and Five Point Differentiator

The general overall effect of the frequency response error is that the instantaneous frequency is lower than desired, and will distort the signal at higher frequencies. The DFT method for approximating a differentiator is the most efficient in terms of frequency response, and an example is provided in figure 2.6, which shows a signal with a large range of instantaneous frequency values.

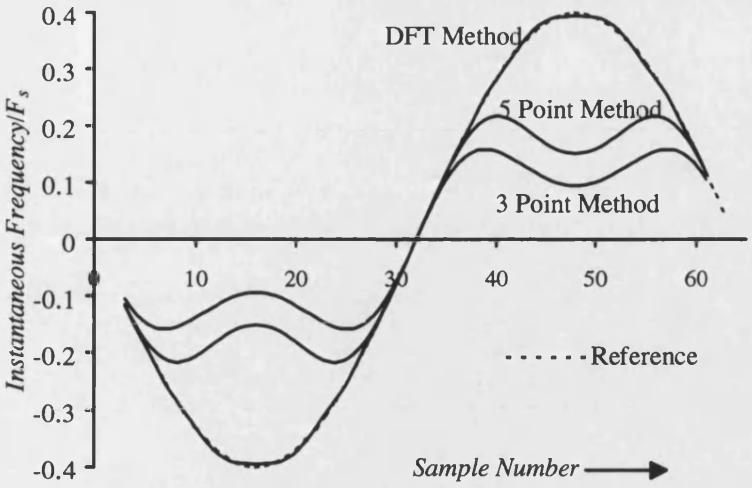


Figure 2.6 : Plots of Instantaneous Frequency against Time for a Sinusoidally Modulated FM Signal

The DFT method does however have drawbacks due to the windowing effects discussed in section 2.3.6. An example is provided in figures 2.7 and 2.8, where the instantaneous frequency is subject to rapid variations.

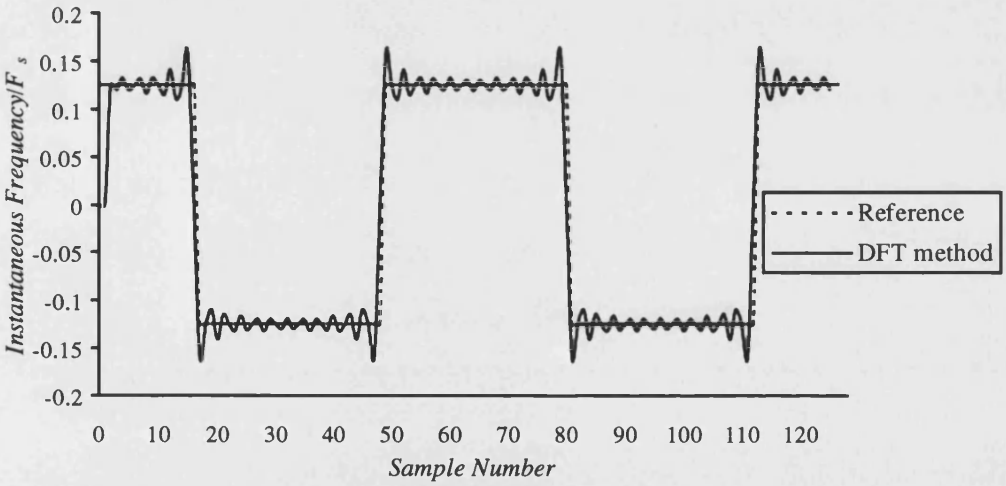


Figure 2.7 : Plots of Instantaneous Frequency Estimate against Time for a Square Wave Modulated FM Signal, DFT Differentiator

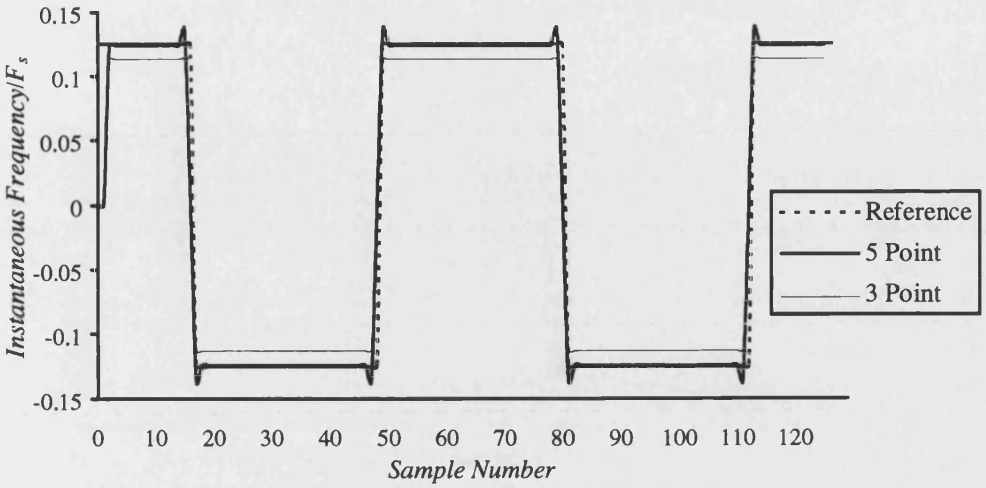


Figure 2.8 : Plots of Instantaneous Frequency Estimate against Time for a Square Wave Modulated FM Signal, FIR Differentiators

From figure 2.7 it is found that the DFT method suffers the worst ringing at the transitions, and the ringing has a low damping factor. In figure 2.8 it is seen that the five point method is also subject to ringing which lasts just one sample either side of the transition. The three point method is not subjected to any ringing.

The analytic approximations therefore require more than a two sample observation frame to provide useful results.

2.12.1.4 Nagy Estimator

This technique was presented in [39] by P.A.J. Nagy. A personal correspondence with the author revealed that the method was derived intuitively. The author attempted to apply a bilinear transform operator to equation (2.22), which provides some useful results, but not a proof. The estimator is given by :

$$f_n(n) = \frac{1}{\pi T_s} \tan^{-1}\left[\mathrm{Im}\left\{ \frac{s(n) - s(n-1)}{s(n) + s(n-1)} \right\} \right] \qquad \text{(H.11)}$$

Following the correspondence, a proof was attempted, and an interpretation was derived based upon a geometric construction. It is shown in the following proof that the estimator approximates the backward finite difference estimator (equation H.2) when the envelope is not subject to fluctuations.

In order to investigate the effects of envelope fluctuations, a signal with combined FM and AM modulation is applied to the estimator. One example is provided in figure 2.9 where a

signal with sinusoidal FM and sinusoidal AM of modulation frequency $0.1F_s$ and AM modulation index of 0.8. It is seen that the instantaneous frequency is subject to large variations due to the AM component.

In general it is found that the fluctuations increase as the true instantaneous frequency and the rate of change of envelope are increased. This example highlights the problems associated with this estimator.

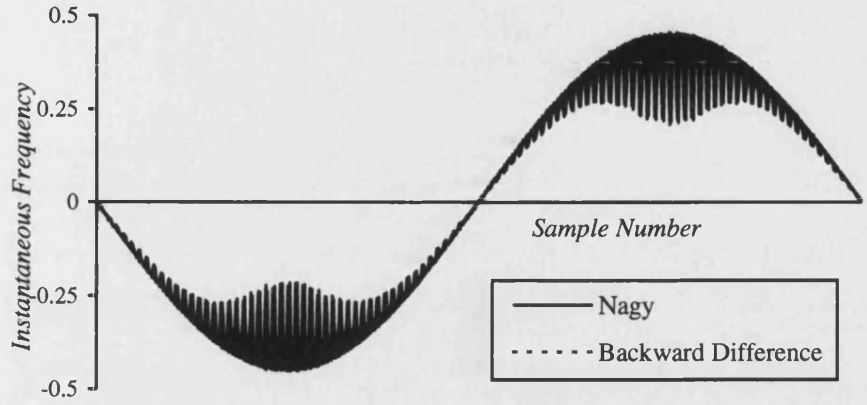


Figure 2.9 : Effect of Amplitude Modulation on the Instantaneous Frequency Operator

Analysis of the Nagy Instantaneous Frequency Estimator

Consider the vectors cast by two analytic signal vectors *Z2* and *Z1* where *Z2* corresponds to $s(n)$ and *Z1* to $s(n\text{-}1)$. The geometric construction in figure 2.10 shows a parallelogram with the two vectors and the angle between the two vectors ψ.

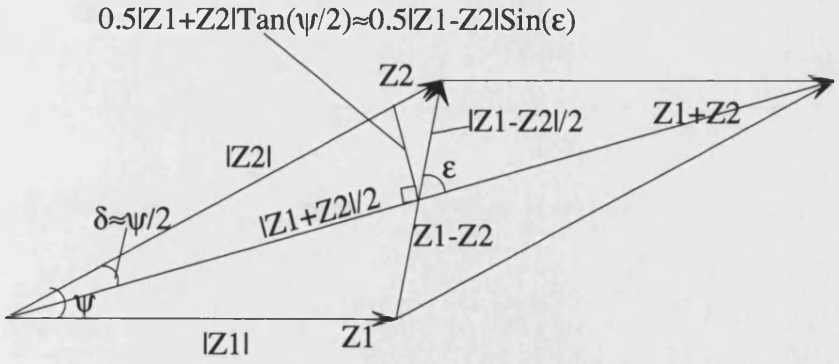


Figure 2.10 : Geometric Construction

It should be noted that if the envelope does not go through a large envelope change between the two samples, the phase δ cast by *Z1+Z2* on *Z2* is approximately given by $\frac{1}{2}\psi$. The

projection from the centre of the parallelogram, perpendicular to *Z1+Z2*, intersecting with *Z2* is expressed in the two forms below :

$$l = \frac{1}{2}|Z1 - Z2|\sin(\varepsilon) \approx \frac{1}{2}|Z1 + Z2|\tan\left(\frac{\psi}{2}\right) \tag{H.12}$$

where ε is the angle between *Z1+Z2* and *Z1-Z2*. This expression may be rearranged to give the following :

$$\tan\left(\frac{\psi}{2}\right) \approx \frac{|Z1 - Z2|}{|Z1 + Z2|}\sin(\varepsilon) \tag{H.13}$$

Note the relationship between the phase difference ψ and the backward difference approximation to instantaneous frequency from equation H.2. The estimate is then expressed by :

$$\begin{aligned}\tan\left(\frac{2\pi f(n)T_s}{2}\right) &\approx \frac{|s(n)-s(n-1)|}{|s(n)+s(n-1)|}\mathrm{Im}\{\exp(j\varepsilon)\} \\ &= \mathrm{Im}\left\{\frac{|s(n)-s(n-1)|}{|s(n)+s(n-1)|}\exp\left(j\left[\arg(s(n)-s(n-1))\right]-\left[\arg(s(n)+s(n-1))\right]\right)\right\}\end{aligned} \tag{H.14}$$

$$\begin{aligned}\tan(\pi f(n)T_s) &\approx \mathrm{Im}\left\{\frac{|s(n)-s(n-1)|\exp\left(j\left[\arg(s(n)-s(n-1))\right]\right)}{|s(n)+s(n-1)|\exp\left(j\left[\arg(s(n)+s(n-1))\right]\right)}\right\} \\ &= \mathrm{Im}\left\{\frac{s(n)-s(n-1)}{s(n)+s(n-1)}\right\}\end{aligned} \tag{H.15}$$

Therefore it follows that :

$$f(n) \approx \frac{1}{\pi T_s}\tan^{-1}\left[\mathrm{Im}\left\{\frac{s(n)-s(n-1)}{s(n)+s(n-1)}\right\}\right] \tag{H.16}$$

This derivation shows that the estimator approximates the backward difference operator when the envelope is approximately constant.

2.12.1.5 Autoregressive Estimate

A method has been presented for instantaneous frequency estimation using autoregressive techniques in two modulation recognition papers generated from the same establishment [37] and [38]. The technique determines the linear prediction coefficients (LPC) and then establishes a peak from the linear prediction spectrum. In order to ensure only one peak, the model order was set to two.

The LPC coefficients a_1 and a_2 are then given by the solution to :

$$\begin{bmatrix} R_0 & R_1 \\ R_0 & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \tag{H.17}$$

where

$$R_k = \sum_{n=0}^{M} x(n)x(n+k) \tag{H.18}$$

and $x(n)$ is the signal which has not been converted into analytic form. Generally this form of equation will be solved using an iterative algorithm, but in this simple form the coefficients may be evaluated through direct inversion of the matrix to give :

$$a_1 = \frac{(R_0 - R_2)R_1}{R_0^2 - R_1^2} \tag{H.19}$$

$$a_2 = \frac{R_0 R_2 - R_1^2}{R_0^2 - R_1^2} \tag{H.20}$$

This produces a frequency domain relationship characterised by the Z transform as :

$$H(Z) = \frac{G}{1 - a_1 Z^{-1} - a_2 Z^{-1}} \tag{H.21}$$

where G is a constant. The instantaneous frequency is related to the roots of the denominator of $H(Z)$ by :

$$f_a(n) = \frac{1}{2\pi T_s} \arg(Z_n) \tag{H.22}$$

and this may be represented in a simple form by :

$$f_a(n) = \frac{1}{2\pi T_s} \tan^{-1}\left(\sqrt{-\left(4a_2 + a_1^2\right)}, a_1\right) \tag{H.23}$$

The smallest window which may be used with this technique is of length three. However, it is found that the estimator does not work well until the window size is greatly increased, e.g. of size thirty.

An example of the estimator is provided for a sinusoidally modulated FM signal in figure 2.11. From the figure it can be seen that the estimator has a poor performance for a 6 sample window, and reasonably bad performance for a 21 sample window.

It is concluded from these results that this is a poor estimator. The poor performance may be due to the number of poles used in the model, but more poles will increase the window length and will also secondly require a root searching function to solve the polynomial, where the dominant peak will have to be selected. It is therefore concluded that this method is unsuitable for modulation recognition.

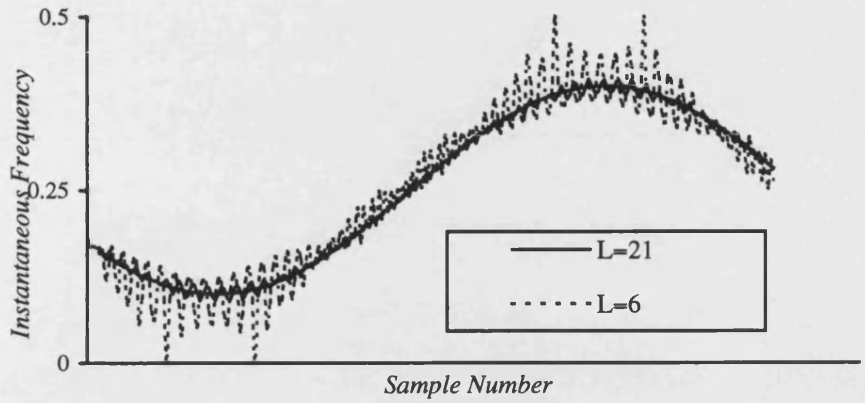


Figure 2.11 : 2-Pole Autoregressive Method of Instantaneous Frequency Estimation

2.12.1.6 Time-Frequency Methods of Estimation

It is possible to use the DFT and the Pseudo Wigner function [42] to estimate the instantaneous frequency of a signal.

The DFT method is often referred to as the Short Time Fourier Transform (STFT), and the minimum window length is two. The Pseudo Wigner distribution has a minimum length of three, but the frequency range is limited to a half of the Nyquist rate, which is half that of the STFT method. The peak of the distribution is often used as a basis for the instantaneous frequency estimate, as it is a simple means of estimation from the distribution.

The Pseudo Wigner function is calculated for frequency bin n and sample number i by :

$$W(i,n) = 2 \sum_{k=-N+1}^{N+1} e^{j2k\pi(n/N)} |h_N(k)|^2 x(i+k)\bar{x}(i-k) \qquad \text{(H.24)}$$

where $h_N(k)$ is the windowing function, and in this case it is unity for the observation period, with a number of zeros to interpolate the spectrum. A technique is given in [42] for evaluation through the Fast Fourier transform (FFT), and the zero padding is therefore set such that the kernel function is a power of two.

The zero padding determines the interpolation which may be present in the frequency estimate, and when the observation window is small, as in this case, the zero padding is essential to reduce quantisation noise in the estimate from the peak search.

The STFT is applied directly through a Fast Fourier transform, and zero padding is equally important in this technique. When evaluation is through the FFT, the magnitude squared component is used for peak searching. In this case the technique is called the Short Time Fast Fourier Transform technique (STFFT).

These techniques are computationally intensive, and their selection will require an improvement in performance over the other simpler techniques which have been described.

2.12.1.7 Comparison of the Techniques in Noise

It is difficult to provide a definitive comparison between the various techniques. A common method looks at the mean squared error of the estimator when a signal is present in Gaussian noise. This method is generally performed on a sinusoid, as the result may be compared with the Cramér-Rao lower bounds [43] [16].

Cramér-Rao lower bounds determine the minimum mean squared error which is theoretically possible to achieve in a given situation. This has been determined for a sinusoid in AGWN of SNR ρ for the case where the amplitude and phase of the signal are unknown [16] as :

$$\mathrm{var}(f) \geq \frac{3}{2\pi^2 T_s^2 \rho N\left(N^2-1\right)} \tag{H.25}$$

where N is the number of independent samples within the observation window. This limit requires further interpretation when applied to instantaneous frequency estimation. The reason for this is that the instantaneous frequency is a circular quantity [6] when applied to discrete observations, which means that the signal is bound within a range.

At low SNR the noise components are bound within a region, and the mean square error tends towards a limiting value. The Cramér-Rao bounds assume that the noise is not bound, and suggest that the error continues to increase as the SNR is decreased. However, at a higher SNR there is no wrapping of the variable, and the comparison against the Cramér-Rao bounds is a valid one. A good estimator will reach the bounds at high SNR, and this is then termed a statistically efficient estimator [12].

2.12.1.7.1 Comparison of the Backward Difference Method With Time-Frequency Techniques

As the Wigner method for estimation requires a minimum of a three point observation window, a three point comparison is made with the STFFT and phase difference techniques.

The phase difference technique has been extended to three points by averaging two adjacent backward difference estimates, and is a special case of the Kay Estimator [12].

The Cramér-Rao bounds for this estimate state that the minimum mean squared error is given by :

$$\mathrm{var}(f) \geq \frac{1}{16\pi^2 T_s^2 \rho} \tag{H.26}$$

Figure 2.12 shows the reciprocal of mean squared error as a function of SNR for a set of simulation runs with unity sampling frequency. From these results it is seen that the error approaches the Cramér-Rao bounds for each estimator between 6 and 8dB SNR. The Wigner and phase difference techniques are seen to outperform the STFFT method in terms of mean squared error.

The Wigner method looks to be the best at first sight, however it should be noted that the range of this method is a quarter of the Nyquist rate as opposed to a half in the other methods. The phase wrapping will therefore occur at a higher SNR than of the other techniques and it is seen that the rate of change of error decreases at around 4dB, which suggests that phase wrapping starts to occur.

The phase differencing technique is seen to be a better method in terms of performance and complexity, and is therefore a more suitable choice than the time-frequency techniques.

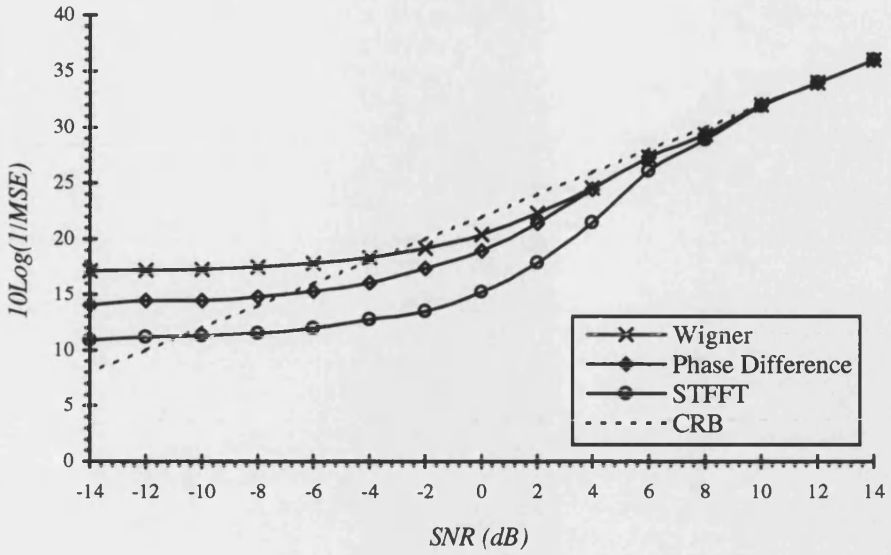


Figure 2.12 : Plots of 1/Mean Squared Error against SNR for a Zero Frequency Sinusoid in Noise

2.12.1.7.2 Comparison of the Time Domain Techniques

The Analytic, Nagy and Backward phase difference techniques are compared using a two sample window. In the Analytic method the derivative is generated through a backward difference technique in order to preserve a two sample window.

The Cramér-Rao bounds for this case are given by :

$$\mathrm{var}(f) \geq \frac{1}{4\pi^2 T_s^2 \rho} \tag{H.27}$$

The results in [2] describe a theoretical development for the mean squared error of the backward difference technique for a sinusoid in Gaussian noise.

Figure 2.13 shows plots of the reciprocal of mean squared error as a function of SNR for the three techniques. It is seen that the Nagy technique performs better in terms of mean squared error than the phase difference technique, and both are better than the Analytic technique.

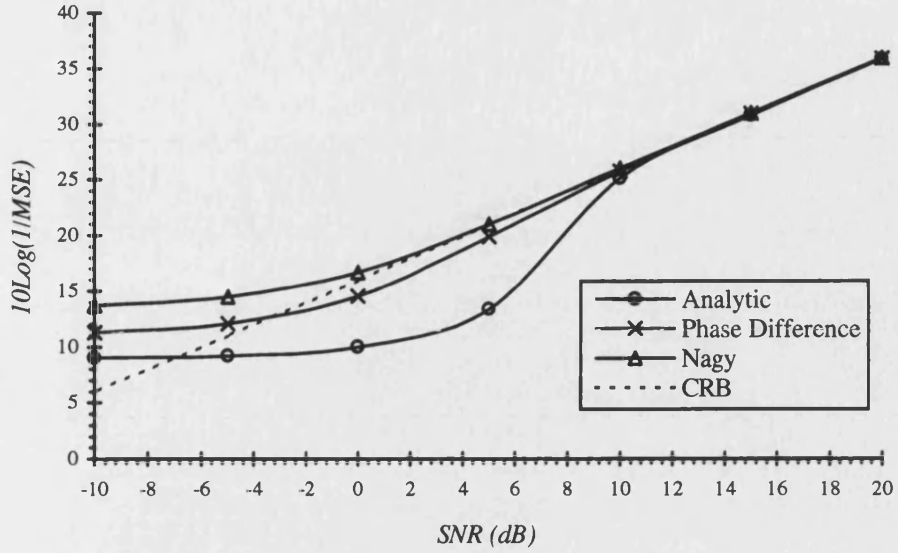


Figure 2.13 : Plots of 1/Mean Squared Error against SNR for a Zero Frequency Sinusoid in Noise

2.12.2 Conclusions

Various instantaneous frequency estimators have been compared, and it has been concluded that a two sample estimator is appropriate for modulation recognition.

The Analytic approximation has been found to yield poor results due to frequency dependent errors. The time-frequency methods are computationally cumbersome and do not perform better in terms of mean squared error in AGWN than that of the phase difference method. The autoregressive method yields poor results for a small window length and is not recommended for use in any method for modulation recognition. The zero crossings method is not useful for an analysis window of two samples.

The Nagy estimator was found to perform well statistically for a sinusoid in noise. However, the effects of envelope fluctuations were seen to distort the frequency estimate.

The Central phase difference technique was found to limit the frequency range of the signal, and consequently the backward phase difference technique was chosen in preference. The statistical performance of this estimator was seen to be good and the computational efficiency is excellent and the estimator was seen to be independent of envelope fluctuations.

The backward difference method for instantaneous frequency estimation is therefore determined to be the best estimator for modulation recognition.

Chapter 3 : The Classification of PSK Signals

3. The Classification of PSK Signals

3.1 Outline

This chapter examines new work on the classification of the number of levels on a PSK signal using decision theoretic methods.

This is an area of research that has received some attention in the published literature [1-7], with most of the work from A. Polydoros et. al. of the University of Southern California and S. Soliman et al. of Qualcomm.

The chapter examines the published work, and further develops some of the techniques in terms of analysis and implementation. The most significant work within this chapter is the development of new methods for PSK classification.

The classifiers discussed from the literature require accurate carrier synchronisation. Included within this work is a new approach to PSK classification which avoids the requirement of carrier synchronisation. These techniques use a differential approach to the classification, which in many ways resembles that of non-coherent differential PSK demodulation. This is found to have a substantial performance degradation associated with it, but is still an attractive technique for PSK classification.

Most of the techniques assume different a-priori information, which will impose a level of attainable performance. This a-priori information has direct impact on the practical implementation of a working system, and in general the information must be estimated before the algorithm may be implemented.

The general direction of the work in this chapter is to develop new techniques which assume ideal conditions, and consequently takes a theoretical bias. Future research is required for the development of these techniques into a practical system, with particular emphasis on the estimation of PSK independent parameters. Some introductory discussion of these issues is provided in Appendix I, and Appendix II looks into the various methods of analytic signal generation which is fundamental in a digitally phase modulated system.

In all carrier coherent classifiers examined in this chapter, the signal is assumed to be a PSK signal of known carrier frequency in AGWN. The symbols are deemed to be equiprobable and of constant envelope. Any deviation from this ideal model is to be modelled as a nuisance factor, and some of the deviations are examined in the following chapter.

3.2 Introduction to the Classifiers

The first decision theoretic PSK classifier apparent in the open literature was presented by A. Polydoros [1], and details a BPSK/QPSK classifier. The classifier statistics were derived from the in-phase and quadrature components of the received signal, and the classifier requires knowledge of the amplitude of the signal as well as the SNR.

The next development was made by S. Soliman [2], who presented a classifier structure which could incorporate a wide range of PSK signals. Classification is based upon the statistical moments of the detected phase. The performance of the technique was improved in subsequent papers [3][4] by improving the accuracy of the model.

Soliman et al. then derived a classifier based on a maximum likelihood statistics for the phase [5], which is more robust and has better classification performance than the Statistical Moments Classifier, but at an increased computational overhead.

Huang and Polydoros improved the classification performance of their original technique [1] by using a power law to discriminate between different PSK types [6]. This was again based upon the in-phase and quadrature channels, and was a binary hypothesis technique. The work was advanced in [7] to incorporate a larger range of PSK classes.

The techniques discussed all require accurate knowledge of the carrier frequency of the signal, and assume that the carrier is removed. Some of the methods require accurate knowledge of the zero phase of the carrier so that the modes of the signal pdf are accurately aligned [2][5][7,cs]. All the methods require knowledge of the SNR of the signal, and those using the in-phase and quadrature signals [1][7] require knowledge of the signal amplitude.

3.2.1 New Classifier Structures Developed Within The Thesis

The work in this thesis develops some new forms of classifiers with different characteristics and assumptions. All of the methods developed have the flexibility to incorporate an arbitrary number of PSK classes within the structure.

The simplest classifier developed is the DFT of Phase Histogram classifier, which has been presented in [8]. This technique does not require knowledge of the carrier zero phase, amplitude or the SNR of the signal, and is a simple method to implement.

The next classifier which has been developed is an extension of the DFT method, and makes use of an SNR estimate to improve the classification performance. This method has been presented in [9], and has a computationally efficient structure.

The globally optimum classifier is derived, which requires knowledge of carrier zero phase, amplitude and the SNR of the signal. The classification performance of the classifier is derived and methods for improving the computational efficiency are also examined. This work is to be presented in [10].

A set of non-coherent classifiers are examined within the thesis and do not require accurate knowledge of the carrier frequency of the signal. These methods are particularly useful, as the carrier frequency is one of the most difficult parameters to determine.

Three non-coherent (asynchronous) techniques have been proposed. The first two techniques perform classification based upon the phase difference modulo 2π, one method using maximum likelihood classification, requiring knowledge of the SNR of the signal, and the other using the DFT of phase difference histogram, which does not require knowledge of the SNR. The third technique performs maximum likelihood classification based upon the phase difference without the modulo 2π arithmetic imposed.

A summary of the various techniques is given in table 3.1, which also includes a summary of the a-priori information each method requires.

	Accurate Carrier Frequency Knowledge	SNR Knowledge	Phase Calculation	Zero Phase Knowledge	Amplitude Knowledge
Existing Techniques					
Statistical Moments	√	√	√	√	X
Optimum Phase	√	√	√	√	X
qLLR method, cs	√	√	X	√	√
qLLR method, ns	√	√	X	X	√
New Techniques					
DFT of Phase Histogram	√	X	√	X	X
Maximum Likelihood DFT of Phase Histogram	√	√	√	X	X
Maximum Likelihood IQ	√	√	X	√	√
Maximum Likelihood Phase Difference Modulo 2π	X	√	√	X	X
Maximum Likelihood Phase Difference non- Modulo 2π	X	√	√	X	X
DFT of Phase Difference Histogram	X	X	√	X	X

Table 3.1 : Description of Classifier Structures

3.3 Chapter Structure

The synchronous phase based classifiers are examined first, starting with the "DFT of Phase Histogram" technique in section 3.4, where the structure and statistical performance of the method are examined in detail. This method is one of the new techniques presented within this thesis, and is then followed in section 3.5 with a maximum likelihood version of the same classifier.

The "Statistical Moments" classifier is discussed in section 3.6, and methods for computational efficiency are discussed with reference to the techniques in chapter 2. Some undesirable effects are attributed to computational inaccuracy, and methods for resolving this are discussed.

The "Optimum Phase" classifier is examined in section 3.7. New techniques are developed for the evaluation of the error performance for a general case and methods are discussed for improving the computational efficiency based upon a cubic spline look-up table.

The synchronous IQ methods follow, starting with the theoretically optimum classifier in section 3.8. This provides the development of the structure and classification performance of the technique and methods are examined for the improvement of the computational efficiency using Padé approximates.

The coherent power law techniques developed by A. Polydoros et al., which are termed qLLR (quasi log likelihood ratio) are examined in section 3.9. The classification performance of these techniques is examined.

Finally the non-coherent methods of PSK classification are developed, where the wrapped phase difference is considered for a maximum likelihood structure in 3.10.1, and a histogram DFT based method in 3.10.2. A further structure is developed in 3.10.3 for the case where the phase difference is not wrapped by 2π.

3.4 DFT of Phase Histogram Classifier

This method has been presented and published in ICASSP '95 [8]. The general technique employed by this method results from an analysis of the pdf of phase of PSK signals at low SNR.

It is noticed in figure 2.2, chapter 2 that at low SNR the pdf of phase for *M*PSK tends towards a sinusoid of *M* cycles. From this observation an algorithm for PSK classification was derived, where the DFT is used to highlight the first Fourier harmonic in the pdf. The peak of the DFT may then be used as a means of classifying *M*. In order to establish the pdf of phase, the pdf is approximated through a histogram, which enables the DFT to be employed effectively.

The performance of the classifier is determined through mathematical analysis assuming that the PSK signal is in AGWN. This analysis places the classifier in a similar category to that of the decision theoretic classifiers.

3.4.1 Development

The performance is developed using the assumption that the signal is modelled effectively by one harmonic at the SNR range of interest and above this, the classification error probability is so low that a multiple harmonic has negligible effect on the classification performance. This assumption is found to be a good one.

The incoming signal is digitally sampled and the phase $\phi(n)$ is determined for L samples.

3.4.2 Histogram Development

The L phase samples $\phi(n)$ are used to build a phase histogram with N bins which approximates the pdf. The following theory characterises the error between the true pdf and this histogram approximation.

The probability of a phase sample entering the i^{th} histogram bin is given by :

$$p_i = \int_{\phi_i - \frac{\Delta}{2}}^{\phi_i + \frac{\Delta}{2}} f(\phi) d\phi \qquad (3.1)$$

where Δ is the bin width. This may be approximated for small bin widths by :

$$p_i \approx \Delta f(\phi_i) \qquad (3.2)$$

As there is a fixed probability of the phase samples entering any histogram bin, the pdf of a bin will be binomially distributed with mean Lp_i and variance :

$$\sigma_i^2 = Lp_i(1-p_i) \tag{3.3}$$

The area of the histogram is $L\Delta$. The histogram is normalised to unity to give a mean μ_i and variance σ_i^2 of :

$$\mu_i = f(\phi_i) \tag{3.4}$$

$$\sigma_i^2 = \frac{1}{L\Delta} f(\phi_i)[1-\Delta f(\phi_i)] \tag{3.5}$$

For small $\Delta f(\phi_i)$ the variance may be approximated by :

$$\sigma_i^2 \approx \frac{1}{L\Delta} f(\phi_i) \tag{3.6}$$

By virtue of the central limit theorem, the errors of all the bins will be normally distributed for large L. The mean variance of the error terms is expressed by :

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \tag{3.7}$$

where $\hat{\sigma}^2$ is the noise variance of the N bin histogram. This then gives :

$$\hat{\sigma}^2 \approx \frac{1}{NL\Delta}\sum_{i=1}^{N} f(\phi_i) \tag{3.8}$$

But $\Delta\sum_{i=1}^{N} f(\phi_i) = 1$, therefore

$$\hat{\sigma}^2 \approx \frac{1}{NL\Delta^2} \tag{3.9}$$

The histogram will have equally spaced points in the interval $[-\pi,\pi]$, therefore $\Delta = \frac{2\pi}{N}$.

Placing this into (3.9), the noise variance is :

$$\hat{\sigma}^2 \approx \frac{N}{4\pi^2 L} \tag{3.10}$$

An example of a 1024 sample, 32 bin histogram for QPSK at 5dB SNR is given in figure 3.1. The underlying sinusoidal nature of the pdf is prominent at this SNR. The pointwise error

between the histogram and the pdf is seen in figure 3.2. This error series is the cause of misclassification, and transpires to the DFT stage.

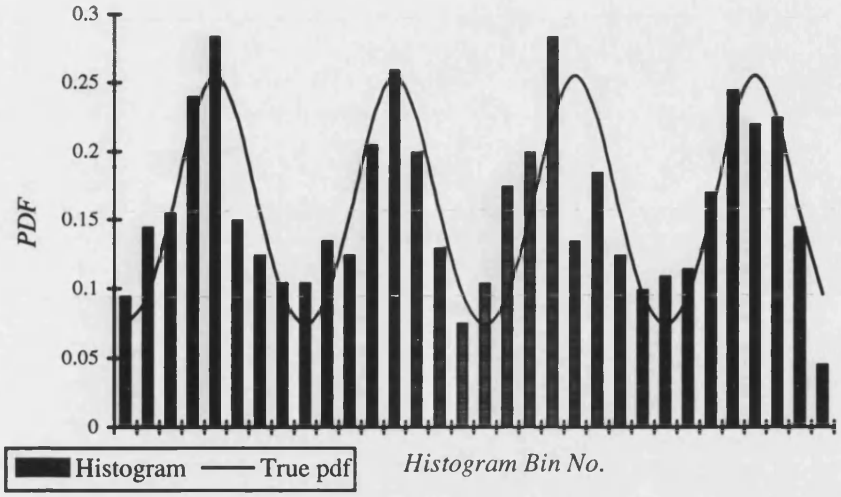


Figure 3.1 : Plots of Histogram and PDF of Phase for QPSK at 5dB SNR, L=1024, N=32

The error signal is modelled in subsequent sections as a zero mean Gaussian series with variance given from (3.10). However, each sample does not have the same variance, and by the nature of a histogram each sample will have some form of dependence. The effect of this will be discussed at a later stage, and the assumption that the noise series is AGWN is used to develop the model.

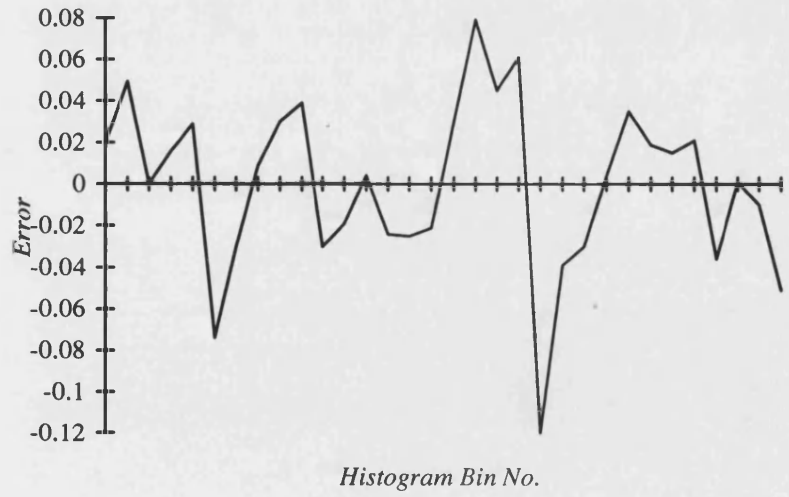


Figure 3.2 : Plot of Histogram Error for QPSK at 5dB SNR, L=1024, N=32

3.4.3 Discrete Fourier Transform

The signal presented to the classifier is modelled as a single sinusoid in the presence of AGWN resulting from the histogram error.

The N points of the histogram are operated on by the discrete time Fourier transform to exploit spectral peaks corresponding to the harmonic terms. From (2.11) it is seen that M level PSK is characterised by a series of spectral lines on the bins which are multiples of M. This indicates that the DFT will be subject to aliasing of harmonics.

When choosing the number of histogram bins N, two conditions must be satisfied. The first is to avoid aliasing of the fundamental harmonic of all potentially present PSK schemes, and the second is to avoid spectral leakage from the harmonics. If M_{MAX} is the highest number of levels on all of the PSK schemes considered, then the two conditions are satisfied by setting N to an integer power of two, and greater than twice M_{MAX}.

The phase of the frequency components is not required in order to characterise the PSK schemes, so magnitude squared of the DFT is used. The magnitude of the DFT is equally valid, but it requires extra processing whilst offering identical performance.

In the previous section it was seen that the histogram is built up from the true pdf plus AGWN. This noise presents a noise floor on the DFT in bins without a harmonic, and will perturb the magnitude of the bins with the harmonic terms. The DFT of the signal $h(n)$ is given by :

$$H(k)=\sum_{n=0}^{N-1} h(n)\exp\left(\frac{-j2\pi kn}{N}\right)=x(k)+jy(k) \tag{3.11}$$

The magnitude squared of this signal is given by :

$$S(k)=x^2(k)+y^2(k) \tag{3.12}$$

In order to normalise the DFT to reflect the magnitude squared of the harmonic components, $S(k)$ is multiplied by the scaling factor $\frac{4}{N^2}$ to give D(k)

$$D(k)=\frac{4}{N^2}S(k) \tag{3.13}$$

From [11] it may be deduced that the noise signal away from the end points has a Chi Squared distribution with two degrees of freedom which is identical to the Rayleigh

distribution. For a noise signal z presented to the DFT with variance σ^2, the distribution is given by :

$$p(z) = \frac{1}{N\sigma^2}\exp\left(-\frac{z}{N\sigma^2}\right) \quad z>0 \tag{3.14}$$

Placing in the expression for the noise variance of the histogram (3.10) :

$$p(z) = \frac{4\pi^2 L}{N^2}\exp\left(-\frac{4\pi^2 Lz}{N^2}\right) \quad z>0 \tag{3.15}$$

It can be shown by a change of variable that the pdf of *D(k)*, *p(y)* is :

$$p(y) = \pi^2 L\exp\left(-\pi^2 Ly\right) \quad y>0 \tag{3.16}$$

where *y* is the random variable of the noise signal.

This result shows that the noise floor is independent of the number of histogram bins *N* and implies that *N* may be made large enough to remove any significant effects of aliasing without affecting the noise floor. It can also be shown (Appendix 3.A) that using different *N* with the same data presents approximately the same result for the same frequency bins, particularly at low *k*. Finally it can be shown that the mean noise floor level is given by :

$$\mu_{floor} = \frac{1}{\pi^2 L} \tag{3.17}$$

When a frequency bin is occupied by a harmonic signal and the histogram noise, the bin is distributed with a non-central Chi-squared distribution, with two degrees of freedom [11]. It can be shown that the distribution *g(x)* of a bin *D(k)* containing a harmonic of amplitude b_m/π and the histogram noise is given by :

$$g(x) = \pi^2 L\exp\left(-L\left[b_m^2 + \pi^2 x\right]\right)\mathrm{I}_0\left(2\pi b_m L\sqrt{x}\right) \quad x>0 \tag{3.18}$$

where $\mathrm{I}_0(z)$ is the modified Bessel function of zero order. It should be noted that this expression is also independent of *N*. Examples of the characteristics of the two pdf's are given in figure 3.3.

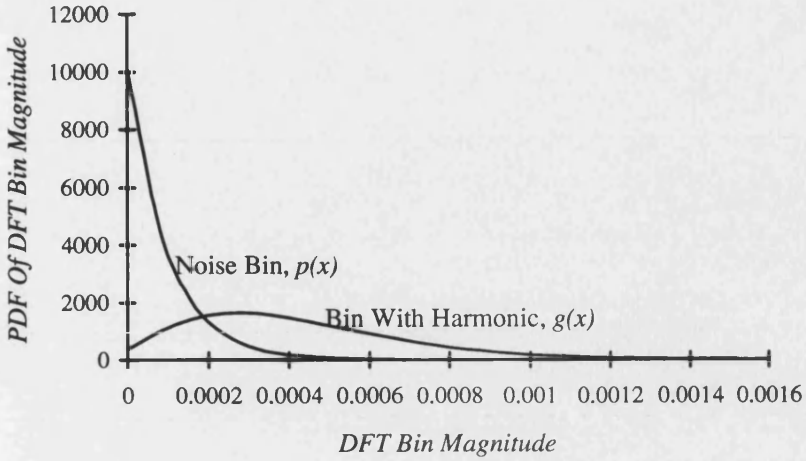


Figure 3.3 : Plots of PDF for the DFT of Phase Histogram for QPSK at 5dB, 1024 Samples

3.4.4 Classification

The classification is achieved by finding the maximum DFT magnitude for the bins which are of interest, $D(\alpha_n)$ where α_n is the number of states in the n^{th} PSK signal. The classified signal is *M*-PSK where :

$$\alpha_M \in MAX[D(\alpha_n)] \tag{3.19}$$

e.g. when 1,2,4 & 8 PSK are to be classified, bins 1,2,4 & 8 of *D(k)* are examined, and if bin 4 is the maximum then the signal is classified as 4 PSK.

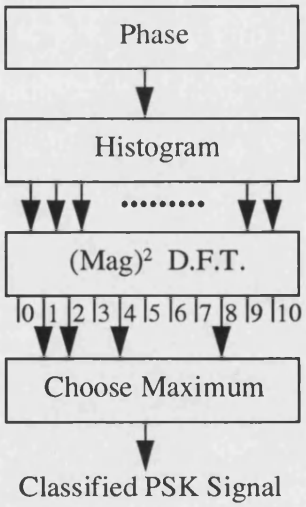


Figure 3.4 : Algorithmic Structure

It can be seen that classification is based on the first harmonic of each PSK signal considered. The higher order harmonics do contain information on the signal class but only the first

harmonics are significant at low SNR, and as the classification procedure does not use the SNR information the higher order harmonics are not included. The overall algorithm is described schematically in figure 3.4.

3.4.5 Probability Of False Classification

Consider the bin containing the signal x with distribution $g(x)$, and n noise bins which are identically and independently distributed with distribution $p(y)$. The probability that the signal lies in the interval $x, x+\delta x$ is given by

$$g(x)\delta x \tag{3.20}$$

Correct classification occurs when all the noise signals are less than x. The probability of correct classification in the interval is therefore :

$$g(x)[1-\Phi(x)]^n\delta x \tag{3.21}$$

where

$$\Phi(x)=\int_x^\infty p(y)dy \tag{3.22}$$

When all of these contributions are summed and in the limit of $\delta x \to 0$, the probability of correct classification is given by :

$$p_{corr}=\int_0^\infty g(x)[1-\Phi(x)]^n dx \tag{3.23}$$

which can be re-written as :

$$p_{corr}=\sum_{i=0}^{n}\frac{n!(-1)^i}{(n-i)!i!}\int_0^\infty g(x)\Phi^i(x)dx \tag{3.24}$$

The probability of error is then given by :

$$p_{err}=1-p_{corr}=\sum_{i=1}^{n}\frac{n!(-1)^{i+1}}{(n-i)!i!}\int_0^\infty g(x)\Phi^i(x)dx \tag{3.25}$$

From (3.16), $\Phi(x)$ is given by :

$$\Phi(x)=\exp\left(-\pi^2 Lx\right) \tag{3.26}$$

Using (3.18) and by manipulating the result in [12] it can be shown that

$$\int_0^\infty g(x)\Phi^i(x)dx = \frac{1}{i+1}\exp\left[-\frac{ib_m^2L}{i+1}\right] \qquad (3.27)$$

Therefore by placing (3.27) in (3.25),

$$p_{err} = \sum_{i=1}^{n}\frac{n!(-1)^{i+1}}{(n-i)!(i+1)!}\exp\left[-\frac{ib_m^2L}{i+1}\right] \qquad (3.28)$$

For a two scheme classification (n=1) :

$$p_{err} = \frac{1}{2}\exp\left[-\frac{b_m^2L}{2}\right] \qquad (3.29)$$

For a four scheme classification (n=3) :

$$p_{err} = \frac{3}{2}\exp\left[-\frac{b_m^2L}{2}\right] - \exp\left[-\frac{2b_m^2L}{3}\right] + \frac{1}{4}\exp\left[-\frac{3b_m^2L}{4}\right] \qquad (3.30)$$

3.4.6 Results

The calculated results are plotted for the DFT classifier (figure 3.5) using the set CW-8PSK for classification and 1024 samples with the SNR ranging from -25dB to 10dB. Also included on the plot in the dashed lines are the simulated results. It can be seen that the simulated results lie close to the theoretical results in the case of CW and BPSK but are slightly below the theoretical results for QPSK and more significantly below for 8PSK.

The error performance is seen to be better than that predicted by the model and becomes significant for QPSK and 8PSK. It is found that there is correlation between histogram error samples which causes the DFT spectrum to become non-uniform.

The result is that at the lower frequency bins the variance is reduced and the higher frequency bins level out to that predicted from the approximations. This is not significant for CW and BPSK, but for QPSK and 8PSK at a higher SNR, the reduction in variance results in an improved performance. However, asymptotically the results tend towards the theory within 0.1dB at 1% misclassification probability. Figure 3.6 shows the error probability for classifying CW-16PSK.

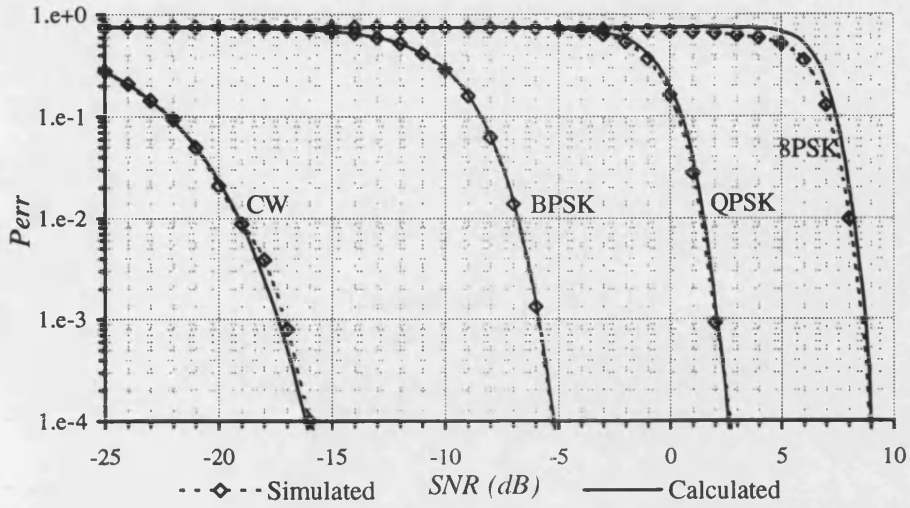


Figure 3.5 : Plots of Misclassification Probability against SNR for CW, BPSK, QPSK and 8PSK, L=1024

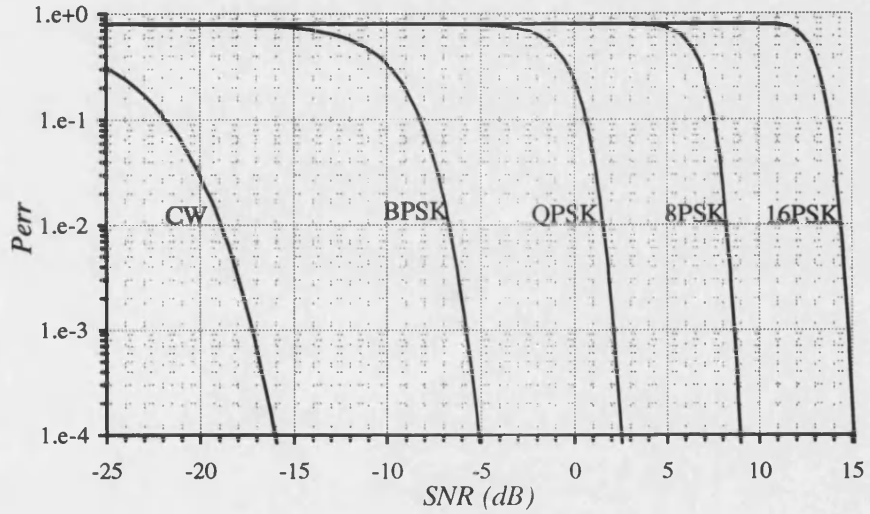


Figure 3.6 : Plots of Misclassification Probability against SNR for CW, BPSK, QPSK, 8PSK, 16PSK, L=1024

3.5 Maximum Likelihood DFT Classifier

The DFT classifier in section 3.4 bases classification on the maximum of the DFT of Phase Histogram, which is independent of the signal SNR. If some SNR information is to be used, then a maximum likelihood structure may be developed. The following development has been presented at GLOBECOM '95 in [9].

3.5.1 Development

When MPSK is transmitted, the pdf of the M^{th} DFT bin may be represented by $g_m(x)$, which was found in section 3.4.3 to be a non-central chi-squared distribution with two degrees of freedom. This is given by :

$$g_m(x) = \pi^2 L \exp\left(-L\left[b_m^2 + \pi^2 x\right]\right) \mathrm{I}_0\left(2\pi b_m L\sqrt{x}\right) \quad x>0 \tag{3.31}$$

Also, the pdf $p(x)$ of the DFT bins without a harmonic component are Rayleigh distributed and are given by :

$$p(x) = \pi^2 L \exp\left(-\pi^2 L x\right) \quad x>0 \tag{3.32}$$

3.5.2 Maximum Likelihood Classifier

In the DFT classifier of section 3.4, classification was based on the maximum of a set of DFT bins, each of which corresponded to a PSK type. This method required no knowledge of the SNR of the signal, but is sub-optimal in the case when the SNR information is known. The technique presented here assumes that the SNR information is available, and uses a maximum likelihood technique to classify the PSK type.

Appendix 3.B shows the general form of maximum likelihood classifier. In this particular problem there is only one set of data to be tested. The probability of the data given the hypothesis of α-PSK transmitted is given from (3.31) and (3.32) by :

$$p(\, x_1, x_2, \ldots x_\alpha, \ldots x_{\max} | \alpha \,\mathrm{PSK}) = \pi^2 L \exp\left(-L\left[b_\alpha^2 + \pi^2 x_\alpha\right]\right) \mathrm{I}_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right) \prod_{\substack{i=1 \\ i\neq\alpha}}^{P} \pi^2 L \exp\left(-\pi^2 L x_i\right) \tag{3.33}$$

where x_i is the output of the i^{th} DFT bin and P is the number of PSK types tested. This may be simplified to :

$$p(\, x_1, x_2, \ldots x_\alpha \ldots . x_{\max} | \alpha \,\mathrm{PSK}) = \exp\left(-Lb_\alpha^2\right) \mathrm{I}_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right) \prod_{i=1}^{P} \pi^2 L \exp\left(-\pi^2 L x_i\right) \tag{3.34}$$

As the maximum is chosen, the terms independent of α may be eliminated, and the classified PSK signal can be represented by :

$$\mathrm{MAX}\left[\exp\left(-Lb_\alpha^2\right) \mathrm{I}_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right), \alpha \in \text{all PSK types}\right] \tag{3.35}$$

increasing function), and applying asymptotic expansions for the Bessel function. For large values, the Bessel function may be represented [13] as:

$$I_0(x) \approx \frac{1}{\sqrt{2\pi x}} \exp(x) \tag{3.36}$$

The likelihood function λ_α may be represented by :

$$\lambda_\alpha = \ln\left[I_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right)\right] - Lb_\alpha^2 \quad x<100 \tag{3.37}$$

which may be evaluated using approximations in [13]. For a large argument the likelihood function is simply expressed by :

$$\lambda_\alpha = Lb_\alpha\left(2\pi\sqrt{x_a} - b_\alpha\right) - \tfrac{1}{2}\ln\left[4\pi^2 b_\alpha L\sqrt{x_\alpha}\right] \quad x>100 \tag{3.38}$$

The overall structure for this classifier is given in figure 3.7.

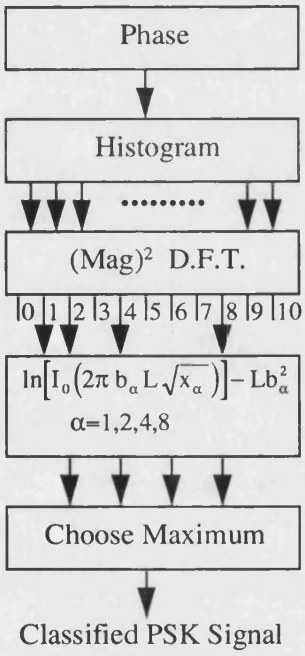


Figure 3.7 : Classifier Structure

3.5.3 Misclassification Probability

The misclassification probability is unfortunately very difficult to calculate. In other similar problems the probability is simpler to calculate due to the a large number of samples being passed through the classifier, thus generating a Gaussian distribution. Therefore the error

performance is evaluated through numerical simulation. Unfortunately this is not the case here, and mathematical evaluation is not practical.

Figures 3.8 and 3.9 show plots of error performance against SNR which have resulted from simulation trials. Figure 3.8 shows the case when L=1024 and CW, BPSK, QPSK and 8PSK are classified. Figure 3.9 shows the case where only BPSK and QPSK are classified.

It can be seen that there is divergence between the QPSK and 8PSK plots of figure 3.8 and the BPSK and QPSK plots of figure 3.9. In a true maximum likelihood classifier it would be found to be coincidental.

These deviations are due to the model deviating from the true pdf at low SNR, and this was observed from the previous DFT classifier (Section 3.4) where the error classification was better than that predicted from the model for QPSK and 8PSK. It was found that the plots converged to the model at a high SNR, and this is also seen in figures 3.8 and 3.9 where the plots converge together, thus indicating a reasonably good maximum likelihood characteristic.

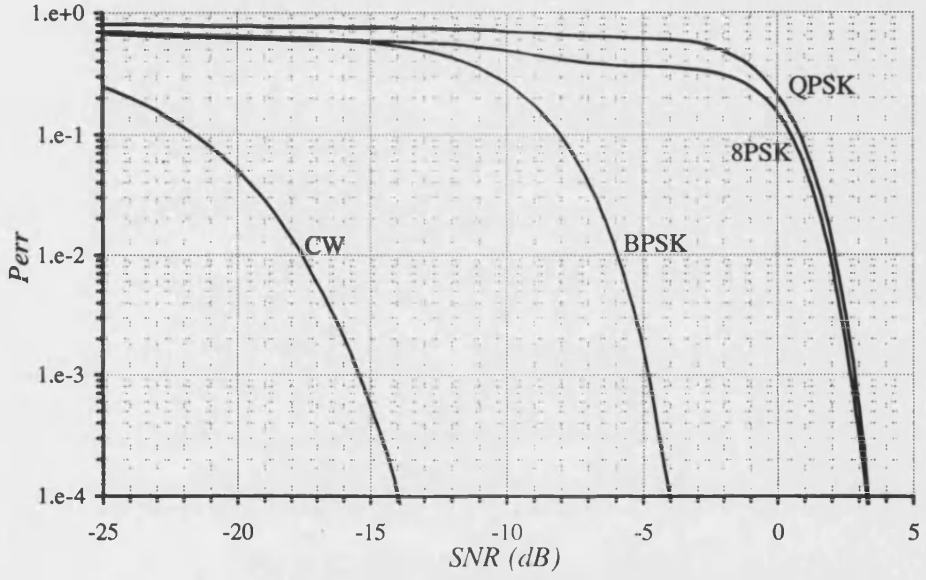


Figure 3.8 : Plots of Misclassification Probability against SNR for CW, BPSK, QPSK and 8PSK, L=1024

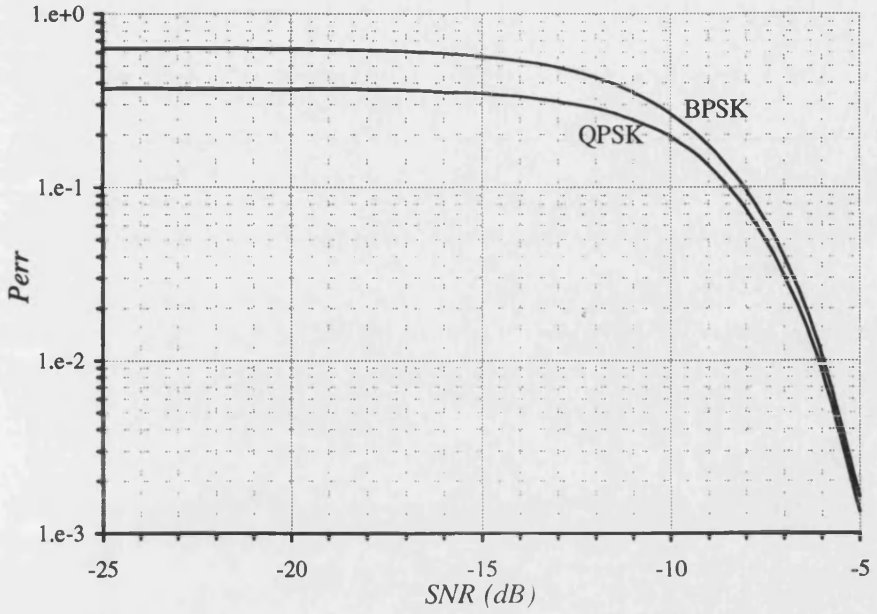


Figure 3.9 : Plots of Misclassification Probability against SNR for BPSK and QPSK, L=1024

3.6 Statistical Moments Classifier

The Statistical Moments Classifier was proposed by Soliman and Hsue in [2]. This form of classifier uses the statistical moment of the phase sample as a feature for classification.

The original approach presented, calculated the theoretical statistical moments through a high SNR approximation. This was then extended to an exact distribution by Yang and Soliman [3], which based the moments on the Fourier series expansion of the phase pdf and was published in journal form in [4].

The moments were then shown to provide discriminating characteristics for an *M*PSK signal, and classification is then based upon a set of calculated thresholds.

3.6.1 Development of Thresholds and Performance

The n^{th} moment of a phase estimate from a discrete time signal ϕ_i is defined as :

$$\hat{m}_n = \frac{1}{L}\sum_{i=1}^{L}\phi_i^n \tag{3.39}$$

The actual moments of an *M* level PSK signal are given by (Appendix 2.F) :

$$m_n = \frac{\pi^n}{n+1} + 2n!\pi^{n-2}\sum_{i=1}^{\infty} b_{iM}\frac{(-1)^i}{(iM)^2}\sum_{k=0}^{\frac{n}{2}-1}\frac{(-1)^k}{(n-2k-1)!(iM\pi)^{2k}} \tag{3.40}$$

It has been found that approximately sixty terms are sufficient for the range of SNR under consideration, and the efficient generation of the Fourier coefficients are detailed in Appendix 2.C.

The expression in (3.39) is the sum of a number of i.i.d. variables, and when L is large the moment estimate becomes a normally distributed variable with mean and variance given by :

$$\mu_M = m_n(M) \tag{3.41}$$

$$\sigma_M^2 = \frac{m_{2n}(M) - m_n^2(M)}{L} \tag{3.42}$$

The thresholds are then considered as a binary hypothesis test between two classes. The hypothesis is given by :

$$\frac{p(\alpha|x)}{p(\beta|x)} = \frac{p(\alpha)}{p(\beta)}\frac{\sigma_\beta}{\sigma_\alpha}\exp\left[-\frac{(x-\mu_\alpha)^2}{2\sigma_\alpha^2} + \frac{(x-\mu_\beta)^2}{2\sigma_\beta^2}\right] \underset{H_\beta}{\overset{H_\alpha}{\gtrless}} 1 \tag{3.43}$$

where x is the measured moment, α and β are the two PSK classes which are tested. Equation (3.43) may be re-arranged to give a threshold of :

$$T = \frac{\mu_\beta\sigma_\alpha^2 - \mu_\alpha\sigma_\beta^2 + \sigma_\alpha\sigma_\beta\sqrt{(\mu_\alpha-\mu_\beta)^2 + (\sigma_\alpha^2-\sigma_\beta^2)\ln\left[\frac{\sigma_\alpha^2}{\sigma_\beta^2}\right]}}{\sigma_\alpha^2 - \sigma_\beta^2} \tag{3.44}$$

This is extended to multi level PSK by assuming that there is negligible influence from non-adjacent PSK types, and the thresholds are then determined from (3.44) as :

$$T_\lambda = \frac{\mu_{\lambda-1}\sigma_\lambda^2 - \mu_\lambda\sigma_{\lambda-1}^2 + \sigma_\lambda\sigma_{\lambda-1}\sqrt{(\mu_\lambda-\mu_{\lambda-1})^2 + (\sigma_\lambda^2-\sigma_{\lambda-1}^2)\ln\left[\frac{\sigma_\lambda^2}{\sigma_{\lambda-1}^2}\right]}}{\sigma_\lambda^2 - \sigma_{\lambda-1}^2} \tag{3.45}$$

where the number of PSK levels is 2^λ and the thresholds are performed for $\lambda>1$. Classification of CW, BPSK, QPSK and 8PSK is then based upon the following thresholds :

$\hat{m}_n < T_1$: CW classified

$T_\lambda \le \hat{m}_n < T_{\lambda+1}$: 2^λ PSK classified

$\hat{m}_n \geq T_{\lambda max}$: $2^{\lambda max}$ PSK classified

The probability of misclassification is then evaluated by examining the cases where the moments variable lies outside the threshold bands. This is evaluated as :

$$p(\text{error}|\lambda\text{PSK Tx}) = Q\left(\frac{T_{\lambda+1} - \mu_\lambda}{\sigma_\lambda}\right) + Q\left(\frac{\mu_\lambda - T_\lambda}{\sigma_\lambda}\right) \tag{3.46}$$

$$p(\text{error}|\text{CW Tx}) = Q\left(\frac{T_1 - \mu_0}{\sigma_0}\right) \tag{3.47}$$

$$p(\text{error}|\lambda \text{ max PSK Tx}) = Q\left(\frac{\mu_{\lambda max} - T_{\lambda max}}{\sigma_{\lambda max}}\right) \tag{3.48}$$

The classification performance is plotted in figures 3.10 and 3.11 for the case where CW, BPSK, QPSK and 8PSK are received with 1024 samples. Figure 3.10 shows the performance of the 8th statistical moment and figure 3.11 shows the 4th moment. From these two plots it can be seen that the performance is very much dependent upon the choice of statistical moment.

A suitable choice of moment is one which is numerically equal to the highest PSK type [2], and any lower than this tends to have a severe performance penalty, which in certain cases can tend towards an error floor.

There are some implementation aspects which must be taken into account when using this technique. The first problem occurs at low SNR, where the variances in the denominator of (3.45) become similar in magnitude. This results in a threshold which is highly sensitive to numerical inaccuracy, and the threshold should be modified to the following approximation :

$$T_\lambda \approx \frac{\mu_{\lambda-1}\sigma_\lambda + \mu_\lambda \sigma_{\lambda-1}}{\sigma_\lambda - \sigma_{\lambda-1}} \tag{3.49}$$

The thresholds must always be within the mean moment boundaries, and a check should be made to ensure that this is the case. In cases where this does not occur, the threshold should be set to the average of the two mean moments.

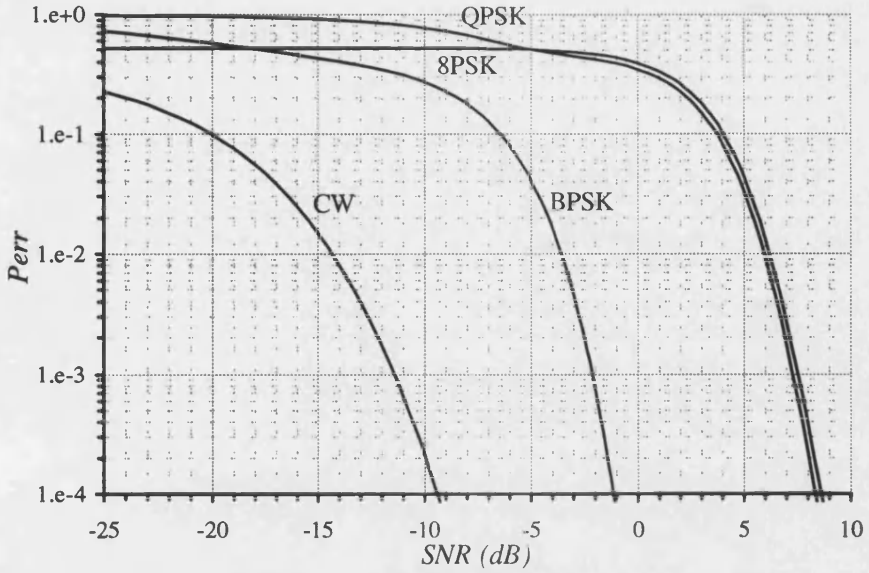


Figure 3.10 : Plots of Misclassification Performance against SNR for CW, BPSK, QPSK & 8PSK, L=1024, 8th Statistical Moment

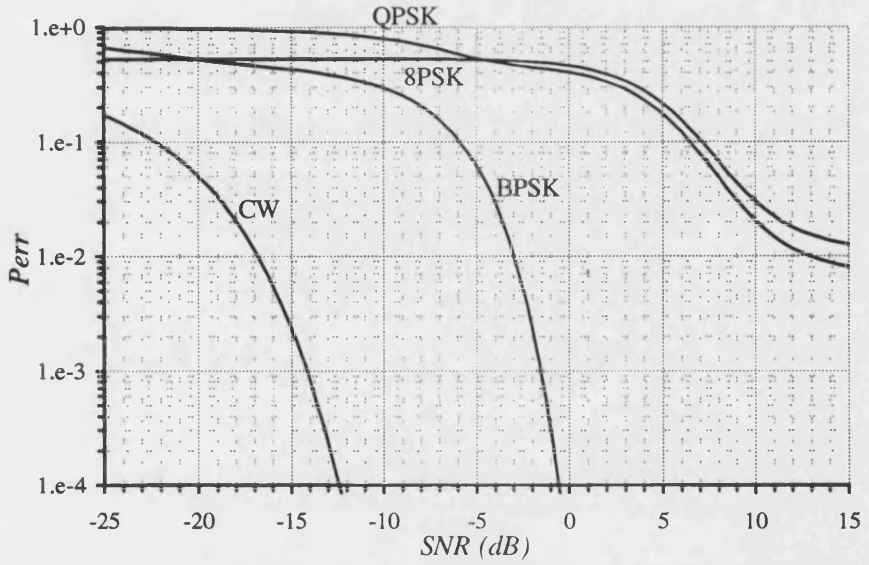


Figure 3.11 : Plots of Misclassification Performance against SNR for CW, BPSK, QPSK & 8PSK, L=1024, 4th Statistical Moment

3.7 Optimum Phase Classifier

The Optimum Phase technique is a maximum likelihood classifier based on the signal phase, and was proposed by Yang and Soliman [5]. The classifier requires knowledge of the carrier zero phase reference in order to operate correctly.

This chapter provides a brief development of the classifier. A technique is developed for evaluating the classification probability which has been published in [14]. One drawback of this technique is the computational burden from the likelihood functions, and methods are examined for improving the computational efficiency.

3.7.1 Summary of the Classifier Structure

The pdf of phase of CW in AGWN was given in Chapter 2, equation 2.5 as :

$$p(\phi)=\frac{1}{2\pi}e^{-\rho}+\frac{1}{2}\sqrt{\frac{\rho}{\pi}}\cos(\phi)e^{-\rho\sin^2(\phi)}\left[1+\mathrm{erf}\left(\sqrt{\rho}\cos(\phi)\right)\right] \tag{3.50}$$

and the pdf of phase of an M level PSK signal is given by :

$$f_M(\phi)=\frac{1}{M}\sum_{n=0}^{M-1}f\left(\phi-\frac{\pi[2n+1]}{M}+\pi\right) \tag{3.51}$$

Alternatively this may be represented in Fourier series form (Chapter 2, equation 2.7) :

$$f(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{m=1}^{\infty}b_m\cos(m\phi) \tag{3.52}$$

The log likelihood function for this form of classifier is detailed from Appendix 3.B as :

$$\lambda_M(\phi)=\ln\left[f_M(\phi)\right] \tag{3.53}$$

Each sample is then passed through a log likelihood function corresponding to each PSK type tested. The results are summed after each sample and the maximum at the end of the trials is the classified signal.

Each sample therefore requires the evaluation of the phase of the signal, followed by an MPSK pdf evaluation and then a logarithm term. This is a computationally intensive task, and simplification techniques may be employed where speed is critical.

Appendix 3.C derives a technique for the efficient evaluation of the log likelihood function based upon a lookup table which is determined at the beginning of each trial. Cubic spline interpolation is then employed, to provide a computationally efficient method of likelihood function generation.

The performance of the classifier is analysed in Appendix 3.D, and the misclassification probability is plotted in figure 3.12 where CW, BPSK, QPSK and 8PSK are assumed potentially present at the receiver.

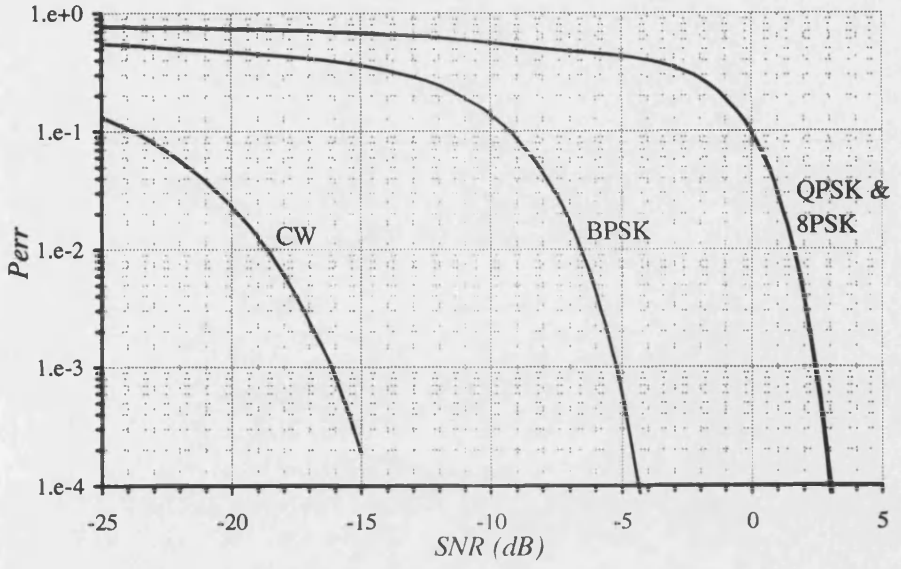


Figure 3.12 : Plots of Misclassification Probability against SNR for CW, BPSK, QPSK and 8PSK, L=1024

3.8 Maximum Likelihood IQ Classifier

This section develops the theoretically optimum classifier. The results of this work are to be published in the MILCOM '96 conference proceedings.

This system performs maximum likelihood classification based on information from the in-phase and quadrature channels directly.

3.8.1 Development

Let x and y be random variables representing the in-phase and quadrature channels respectively. The channel is assumed to be perturbed by AGWN, and the joint pdf of x and y when a CW is transmitted with phase ψ is given by :

$$p(x,y)=\frac{1}{2\pi\sigma^2}\exp\left[-\frac{\left(x-A\cos(\psi)\right)^2+\left(y-A\sin(\psi)\right)^2}{2\sigma^2}\right] \tag{3.54}$$

Expanding this expression :

$$p(x,y)=\frac{1}{2\pi\sigma^2}\exp\left[-\frac{x^2+y^2+A^2-2A\left(x\cos(\psi)+y\sin(\psi)\right)}{2\sigma^2}\right] \tag{3.55}$$

For *M*PSK the pdf consists of the scaled sum of phase shifted versions of the CW pdf (Chapter 2.3.2), given by :

$$p_M(x,y)=\frac{1}{2\pi\sigma^2 M}\exp\left[-\frac{x^2+y^2+A^2}{2\sigma^2}\right]\sum_{j=0}^{M-1}\exp\left[\frac{A}{\sigma^2}\left(x\cos(\theta_j)+y\sin(\theta_j)\right)\right] \tag{3.56}$$

where θ_j are the phase states of the *M*PSK signal and :

$$\theta_j=\pi\frac{2j+1-M}{M} \tag{3.57}$$

Using the maximum likelihood classifier of Appendix 3.B, the log likelihood function is given by :

$$\lambda_M(x,y)=\ln\left\{\frac{1}{2\pi\sigma^2 M}\exp\left[-\frac{x^2+y^2+A^2}{2\sigma^2}\right]\sum_{j=0}^{M-1}\exp\left[\frac{A}{\sigma^2}\left(x\cos(\theta_j)+y\sin(\theta_j)\right)\right]\right\} \tag{3.58}$$

As the maximum statistic is used, the multiplicative terms independent of M may be removed to leave :

$$\lambda_M(x,y)=\ln\left\{\frac{1}{M}\sum_{j=0}^{M-1}\exp\left[\frac{A}{\sigma^2}\left(x\cos(\theta_j)+y\sin(\theta_j)\right)\right]\right\} \tag{3.59}$$

This may be simplified for different forms of PSK signals (Appendix 3.E) to give :

$$\lambda_1(x,y)=\frac{A}{\sigma^2}x \tag{3.60}$$

$$\lambda_2(x,y)=\ln\left[\cosh\left(\frac{A}{\sigma^2}y\right)\right] \tag{3.61}$$

$$\lambda_4(x,y)=\ln\left[\cosh\left(\frac{A}{\sqrt{2}\sigma^2}x\right)\right]+\ln\left[\cosh\left(\frac{A}{\sqrt{2}\sigma^2}y\right)\right] \tag{3.62}$$

$$\lambda_8(x,y)=\ln\left[\frac{1}{2}\{\cosh(ax)\cosh(by)+\cosh(bx)\cosh(ay)\}\right] \tag{3.63}$$

where $a=\sqrt{\frac{\sqrt{2}+1}{2\sqrt{2}}}$ and $b=\sqrt{\frac{\sqrt{2}-1}{2\sqrt{2}}}$

Each sample set is then placed through the likelihood function for each PSK type assumed present. This produces the likelihood output ℓ_j for j level PSK. Classification is then based upon the maximum likelihood function at the end of a trial, i.e. :

$$\ell_j=\sum_{i=1}^{L}\lambda_j(x_i,y_i) \tag{3.64}$$

and the classified PSK signal M is :

$$M = \text{MAX}\left[\ell_j, j\right]_{j \in \text{All PSK Types}} \tag{3.65}$$

The general structure of this technique is given in figure 3.13. This structure enables any number of PSK signals to be classified.

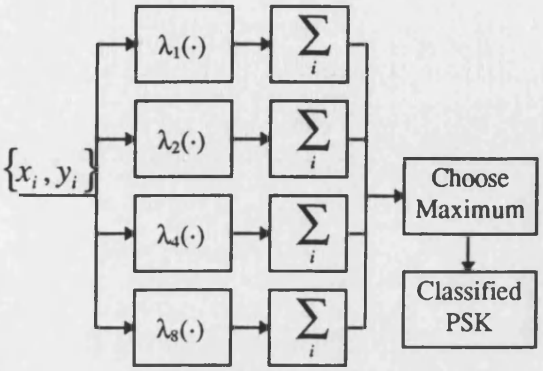


Figure 3.13 : Maximum Likelihood Classifier Structure for PSK

It is seen from (3.60) to (3.63) that this method requires knowledge of the noise variance and the amplitude of the signal. The noise variance is related to the SNR of a signal by :

$$\sigma^2 = \frac{A^2}{2\rho} \tag{3.66}$$

3.8.2 Classification Performance

The classification performance is derived for the case where BPSK and QPSK are discriminated, which provides a relatively simple solution. The expression for ℓ_M in (3.64) consists of the sum of a number of independent variables. If L is large enough, ℓ_M will tend towards a normally distributed variable by virtue of the central limit theorem. The mean and variance of the functions are derived for each likelihood function given a particular PSK type transmitted. The correlation between the distributions is determined, from which the probability of false classification may be obtained.

The statistics of the likelihood functions result from the statistics of x and y passed through the likelihood functions. In the case of BPSK there is one variable, and using an extension of [15] the n^{th} moment is given by :

$$m(n) = \int_{-\infty}^{\infty} g^n(y) f_y(y) dy \tag{3.67}$$

For QPSK the expression is given in a more general form as :

$$m(n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g^n(x,y) f_{xy}(x,y)\,dx\,dy \tag{3.68}$$

This is a two dimensional integral which may be simplified since the expression for λ_4 in (3.62) may be separated into two independent parts as :

$$\lambda_4(x_i, y_i) = \ln\left[\cosh\left(\sqrt{\rho}\,x_i\right)\right] + \ln\left[\cosh\left(\sqrt{\rho}\,y_i\right)\right] = g_\alpha(x) + g_\beta(y) \tag{3.69}$$

(3.68) may now be expressed as :

$$m(n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[g_\alpha(x) + g_\beta(y)\right]^n f_x(x) f_y(y)\,dx\,dy \tag{3.70}$$

For the first moment, this simplifies to :

$$m(1) = \int_{-\infty}^{\infty} g_\alpha(x) f_x(x)\,dx + \int_{-\infty}^{\infty} g_\beta(y) f_y(y)\,dy \tag{3.71}$$

and the second moment :

$$m(2) = \int_{-\infty}^{\infty} g_\alpha^2(x) f_x(x)\,dx + \int_{-\infty}^{\infty} g_\beta^2(y) f_y(y)\,dy + 2\int_{-\infty}^{\infty} g_\alpha(x) f_x(x)\,dx \int_{-\infty}^{\infty} g_\beta(y) f_y(y)\,dy \tag{3.72}$$

These expressions cannot be evaluated analytically, but may be evaluated using numerical techniques to find the mean μ_{ij} and variance σ_{ij}^2, where j level PSK is transmitted and is tested against the i level PSK statistic.

The output statistics of ℓ_2' and ℓ_4' are not independent, and their correlation coefficient r_j is given by [15] as :

$$r_j = \frac{E[\lambda_2\lambda_4] - \mu_{2,j}\mu_{4,j}}{\sigma_{2,j}\sigma_{4,j}} \tag{3.73}$$

Using the definitions in (3.69) and (3.61) and the fact that x and y are independent, the mean of the cross product terms may be expressed as :

$$E[\lambda_2\lambda_4] = \int_{-\infty}^{\infty} \lambda_2(y) g_\beta(y) f_y(y)\,dy + \int_{-\infty}^{\infty} \lambda_2(y) f_y(y)\,dy \int_{-\infty}^{\infty} g_\alpha(x) f_x(x)\,dx \tag{3.74}$$

As it is assumed that L is large, the central limit theorem states that the statistics of the likelihood function may be approximated by normally distributed variables, with mean and variance given by :

$$\mu o_{i,j} = L\mu_{i,j} \tag{3.75}$$

$$\sigma o_{i,j}^2 = L\sigma_{i,j}^2 \tag{3.76}$$

3.8.2.1 Error Probability given BPSK transmitted

For BPSK, the x and y statistics used to evaluate the likelihood function mean, variance and correlation coefficient are given by :

$$f_x(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{3.77}$$

$$f_y(y) = \frac{1}{\sqrt{2\pi}} e^{-\rho} \exp\left(-\frac{y^2}{2}\right) \cosh\left(\sqrt{2\rho}.y\right) \tag{3.78}$$

An error occurs when $(\lambda_2' - \lambda_4') < 0$. The statistics of this difference are normally distributed with mean and variance given by :

$$\mu d = \mu o_{2,2} - \mu o_{4,2} \tag{3.79}$$

$$\sigma d^2 = \sigma o_{2,2}^2 + \sigma o_{4,2}^2 - 2r_2 \sigma o_{2,2} \sigma o_{4,2} \tag{3.80}$$

Finally, the probability of error is simply expressed as :

$$p(error|BPSK) = \Phi\left(\frac{-\mu d}{\sigma d}\right) = \Phi\left(\frac{-\sqrt{L}\left[\mu_{2,2} - \mu_{4,2}\right]}{\sqrt{\sigma_{2,2}^2 + \sigma_{4,2}^2 - 2r_2 \sigma_{2,2} \sigma_{4,2}}}\right) \tag{3.81}$$

where $\Phi(x)$ is the cumulative normal distribution function, and is given by :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt \tag{3.82}$$

3.8.2.2 Error Probability given QPSK transmitted

When QPSK is transmitted the classification error statistics are derived in a similar way, using :

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\rho}{2}} \exp\left(-\frac{x^2}{2}\right) \cosh\left(\sqrt{\rho}.x\right) \tag{3.83}$$

$$f_y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\rho}{2}} \exp\left(-\frac{y^2}{2}\right) \cosh\left(\sqrt{\rho}.y\right) \tag{3.84}$$

and the probability of error is expressed as :

$$p(error|QPSK) = \Phi\left(\frac{-\sqrt{L}\left[\mu_{4,4} - \mu_{2,4}\right]}{\sqrt{\sigma^2_{4,4} + \sigma^2_{2,4} - 2r_4\sigma_{4,4}\sigma_{2,4}}}\right) \tag{3.85}$$

The classification performance resulting from the above theory is provided in figure 3.14 for the BPSK/ QPSK case and is verified by the results from simulation. This is extended in figure 3.15 to the case where CW, BPSK, QPSK and 8PSK are potentially transmitted.

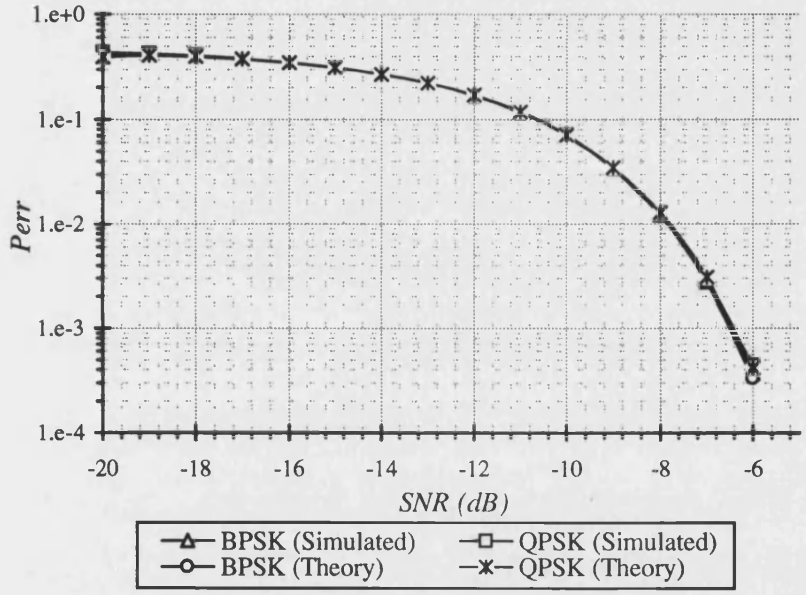


Figure 3.14 : Plots of Misclassification Probability against SNR for CW & BPSK, L=1024

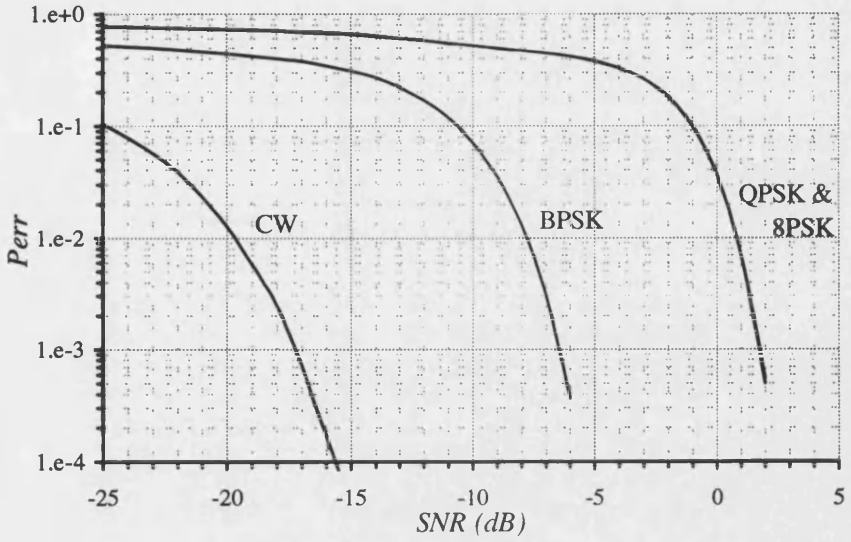


Figure 3.15 : Plots of Misclassification Probability against SNR for CW, BPSK, QPSK and 8PSK, L=1024

3.8.3 Improvement in Computational Efficiency

The computation of the likelihood function requires the evaluation of a set of functions on each sample pair, which will cause the bulk of the computational complexity when a large number of samples is employed. We therefore examine these functions carefully, to see how the computational complexity can be reduced.

The evaluations of the BPSK and QPSK likelihood functions are dominated by the evaluation of a function in the form of $\ln(\cosh(x))$. For a large index, this may be expressed simply as :

$$\ln(\cosh(x)) = x - \ln(2) \tag{3.86}$$

There are a number of low index expressions which may be applied, with a trade off between complexity and error. Each of these will have an optimum threshold point where the function error is equal to that of (36), where the decision is made to the high or low SNR approximation. Two examples are given using Padé approximations. The first results in a peak error of 0.4% :

$$\ln(\cosh(x)) = \frac{\frac{1}{2}x^2 + \frac{19}{124}x^4 + \frac{211}{39060}x^6}{1 + \frac{44}{93}x^2 + \frac{885}{19530}x^4} \qquad x<2.44 \tag{3.87}$$

A simpler expression which results in a peak error of 1.8% is :

$$\ln(\cosh(x)) = \frac{\frac{1}{2}x^2 + \frac{1}{20}x^4}{1+\frac{12}{45}x^2} \quad x<1.9 \tag{3.88}$$

3.9 Coherent Power Law Classifier (qLLR)

This method was introduced by C. Huang and Polydoros in [7]. The technique is a simple M^{th} power law method. This is a common method used within communication systems to establish the carrier frequency of a digital signal, but there is a slightly different approach taken with this method. In this method the carrier frequency of the signal is assumed to be known.

Two cases are examined, the carrier phase synchronous case (cs) and the phase non-synchronous case (ns). Both methods assume that the carrier frequency is accurately known, and the cs method assumes that the carrier reference phase is also known.

The ns statistic looks at the modulus of the transformed phase vector in order to determine a mean component, i.e.

$$qns(M) = \left| \sum_{i=1}^{N} s(i)^M \right| \tag{3.89}$$

The cs case uses the fact that the mean transformed signal will lie on the negative real axis when the power is the PSK type or greater. The statistic used for this is :

$$qcs(M) = \sum_{i=1}^{N} \mathrm{Re}\left[s(n)^M \right] \tag{3.90}$$

The general procedure is to consider the problem as a binary hypothesis test, and to analyse the test statistic for adjacent PSK classes. This is performed in ascending order, where a decision is made at each stage. If the statistic is deemed to have a mean component, the test stops and the PSK type is classified according to the power used at this point. If it is not then the next highest power is applied until the highest PSK type is classified.

In the example given in figure 3.16, CW, BPSK, QPSK and 8PSK are considered potentially present at the receiver. Each PSK type is tested in turn, starting with CW. If the statistic is greater than a threshold then CW is classified, otherwise the test is performed for a power two and the test occurs for BPSK, and so on. If QPSK is not positive then 8PSK is classified.

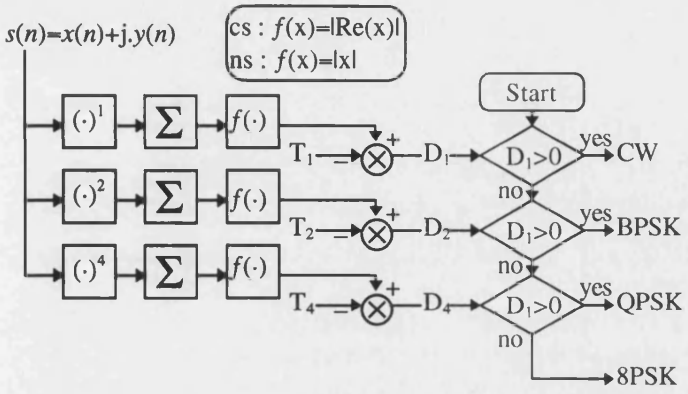


Figure 3.16 : qLLR Classifier for CW, BPSK, QPSK & 8PSK Classification

The technique in [7] scales the signal at the output of the integrate and dump such that the noise variance is unity, and the signal amplitude is $\sqrt{2\rho}$ where ρ is the signal to noise ratio.

3.9.1 Carrier Phase Synchronous Threshold

It is found [7] that the noise variance at the output of the cs test statistic has a mean value of $-N\sqrt{2\rho}$ when the power is the same as the PSK order, and zero when the PSK order is greater than the power. The distribution of the statistic is of a Gaussian form due to the central limit theorem and the variance of the noise is approximately equal in both cases. The threshold is therefore set at :

$$Tcs(M) = \frac{N}{2}(2\rho)^{\frac{M}{2}} \tag{3.91}$$

Figure 3.17 shows plots of the classification performance for this technique for the cases of CW, BPSK, QPSK and 8PSK being potentially transmitted.

3.9.2 Carrier Phase Non-Synchronous Threshold

The distribution of the non-carrier synchronous statistic is of a non-central chi squared form of two degrees of freedom. When the power is less than that of the PSK type, the pdf is represented by a Rayleigh distribution. The variance is again assumed approximately equal in both cases and the threshold is determined by the intersection of the two pdf's to give :

$$Tns(M) \approx (2\rho)^{\frac{M}{2}} V_M \, I_0^{-1}\left[\exp\left(\frac{N}{2V_M}\right)\right] \tag{3.92}$$

where :

$$V_M = \sum_{i=0}^{M} \frac{(M!)^2}{2(i!)[(M-i)!]^2 \rho^i} \tag{3.93}$$

Results of the simulated classification performance are provided in figure 3.18 for the same conditions as the cs case. A simplification has been proposed in [7] to apply the cs threshold to the ns case and figure 3.19 plots the simulated results from this simplified threshold. It can be seen that the plots are subject to sub-optimal performance. QPSK and BPSK reach a local minima at -6 and -17dB respectively, but eventually converge to a negative gradient above -2 and -13dB respectively. The performance of 8PSK is poor for an SNR <0dB, but when the misclassification probability < 10^{-2}, the performance is seen to be close to that of the complete threshold.

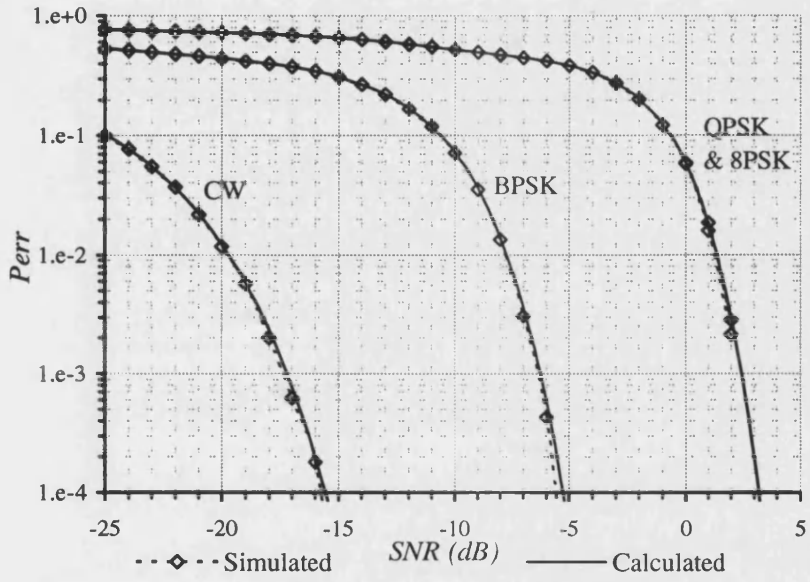


Figure 3.17 : Plots of Misclassification Performance against SNR for CW, BPSK, QPSK & 8PSK, L=1024, Carrier Phase Synchronous

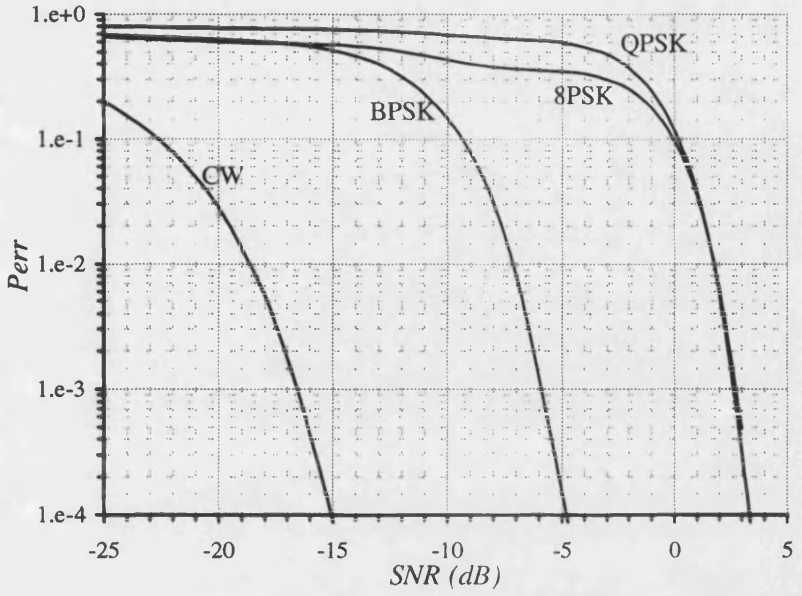


Figure 3.18 : Plots of Misclassification Performance against SNR for CW, BPSK, QPSK & 8PSK, L=1024, Carrier Phase Non-Synchronous

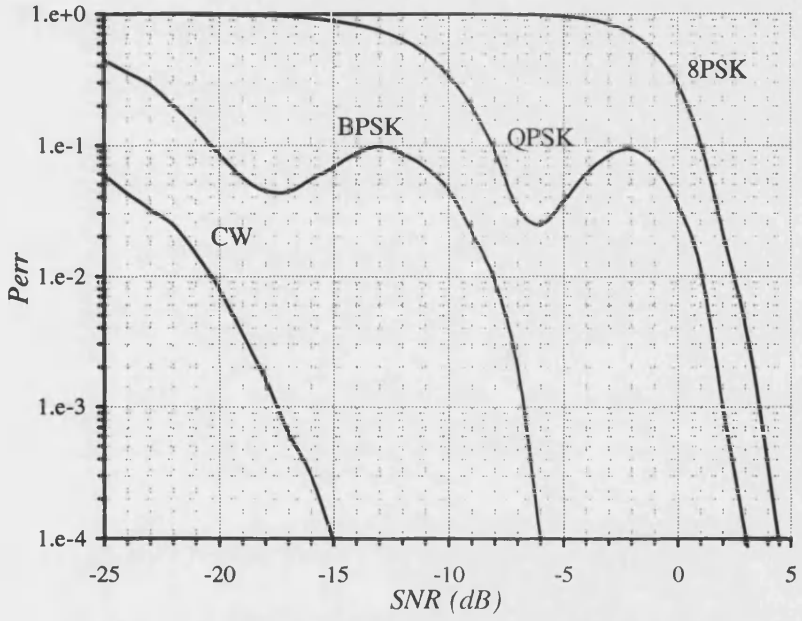


Figure 3.19 : Plots of Misclassification Performance against SNR for CW, BPSK, QPSK & 8PSK, L=1024, Carrier Phase Non-Synchronous, Simplified Threshold

3.10 Non Coherent Techniques

The techniques in the previous section make the assumption that the carrier frequency of the incoming signal is accurately known. The following work develops a range of classifiers which do not require accurate knowledge of the carrier frequency, and are therefore referred to as non-coherent techniques. The published literature does not address this form of decision theoretic techniques for PSK, and all the work is of a novel aspect.

The phase difference of the PSK symbols is used for the classification, and requires knowledge of the symbol timing and rate. This is a desirable parameter to have in any PSK classifier as it enables noise reduction through a matched filter/ integrate and dump section. The symbol timing can however be estimated without coherent knowledge of the carrier frequency, and such techniques.

The technique of using the phase difference of PSK signals is a common technique in non-coherent PSK demodulation and is known as differential PSK. The following development derives a set of PSK classifier structures in the presence of AGWN.

There are three classifiers developed. The first two use the phase difference of the signal modulo 2π, using a maximum likelihood and a DFT structure. The third uses a maximum likelihood structure on the phase difference without modulo 2π.

3.10.1 Maximum Likelihood Phase Difference (Modulo 2π) Classifier

Let ψ be the phase difference of the signal modulo 2π of two adjacent symbols at the output of the matched filter. The relationship may be written as :

$$\psi_i = \arg[s_i \bar{s}_{i-1}] \tag{3.94}$$

where i is the symbol number, and s_i is the i^{th} complex output of the matched filter. Assuming that each sample is statistically independent the pdf of phase difference ψ is described in (Chapter 2, equation 2.29) :

$$p_M(\psi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}(b_{n.M})^2 \cos(nM\psi) \tag{3.95}$$

The graph in figure 3.20 compares the pdf of phase and pdf of phase difference for QPSK at 10dB SNR. It can be seen from the graph that the pdf of phase difference is less well defined, which suggests that there will be an inherent performance loss.

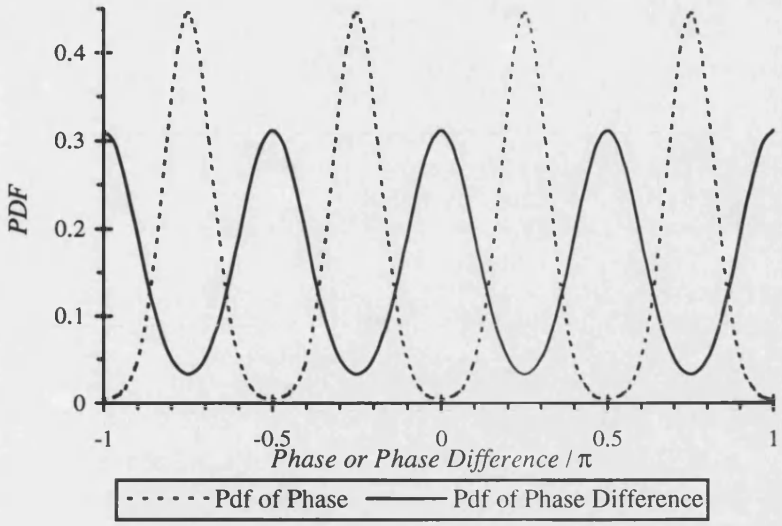


Figure 3.20 : Plots of PDF of Phase and PDF of Phase Difference for QPSK at 10dB SNR

3.10.1.1 Maximum Likelihood Classifier

The log likelihood function for a maximum likelihood classifier is given by (Appendix 3.B) :

$$\lambda_M(\psi) = \ln\left[p_m(\psi)\right] \tag{3.96}$$

Removing the constant terms, this is evaluated as :

$$\lambda_M(\psi) = \ln\left[1 + 2\sum_{n=1}^{\infty}\left(b_{n.M}\right)^2 \cos(nM\psi)\right] \tag{3.97}$$

Each likelihood function is applied to each sample in turn to produce a result for each PSK type considered. This is represented for j PSK with L samples as :

$$\lambda_j = \sum_{i=1}^{L}\lambda(\psi_i) \tag{3.98}$$

This classifier can obviously not be evaluated for an infinite number of Fourier harmonic terms, and the series must be truncated. A single term classifier is investigated, which provides a simple likelihood function, i.e. :

$$\lambda_M(\psi) = \ln\left[1 + 2\left(b_M\right)^2 \cos(M\psi)\right] \tag{3.99}$$

3.10.1.2 Simplification

The expressions for b_2 and b_4 may be evaluated directly from the results in Chapter 2, Appendix 2.B.

$$b_2 = \frac{\rho + e^{-\rho} - 1}{\rho} \tag{3.100}$$

$$b_4 = \frac{\rho^2 - 4\rho + 6 - 2e^{-\rho}(\rho + 3)}{\rho^2} \tag{3.101}$$

The computational burden associated with the generation of the cosine terms may be greatly reduced using the iterative techniques described in Chapter 2, Appendix 2.C and Appendix 2.D which requires one sine and cosine generation followed by a set of simple recurrences.

The evaluation of the first cosine term may be achieved by observing the trigonometric identity :

$$\cos(2\psi) = \frac{1 - \tan^2(\psi)}{1 + \tan^2(\psi)} \tag{3.102}$$

From (3.94) it is seen that :

$$\tan(\psi_i) = \frac{\Im(s_i \bar{s}_{i-1})}{\Re(s_i \bar{s}_{i-1})} \tag{3.103}$$

Therefore :

$$\cos(2\psi) = \frac{\left[\Re(s_i \bar{s}_{i-1})\right]^2 - \left[\Im(s_i \bar{s}_{i-1})\right]^2}{\left[\Re(s_i \bar{s}_{i-1})\right]^2 + \left[\Im(s_i \bar{s}_{i-1})\right]^2} \tag{3.104}$$

If a high order PSK signal is to be classified, the iterative algorithm will use :

$$\sin(2\psi) = \frac{\Re(s_i \bar{s}_{i-1})\Im(s_i \bar{s}_{i-1})}{\left[\Re(s_i \bar{s}_{i-1})\right]^2 + \left[\Im(s_i \bar{s}_{i-1})\right]^2} \tag{3.105}$$

3.10.1.3 Evaluation of Classification Performance

The classification performance may be determined in much the same way as for the Optimum Phase classifier of section 3.7. This has been obtained for the BPSK/QPSK case for a one term and an eighty term classifier, and is shown in figure 3.21.

It can be seen that the one term classifier has a performance which is as good as that of the eighty term classifier, which indicates that the approximation is an extremely good one. Simulation points are also included on the graph, and verify the theory.

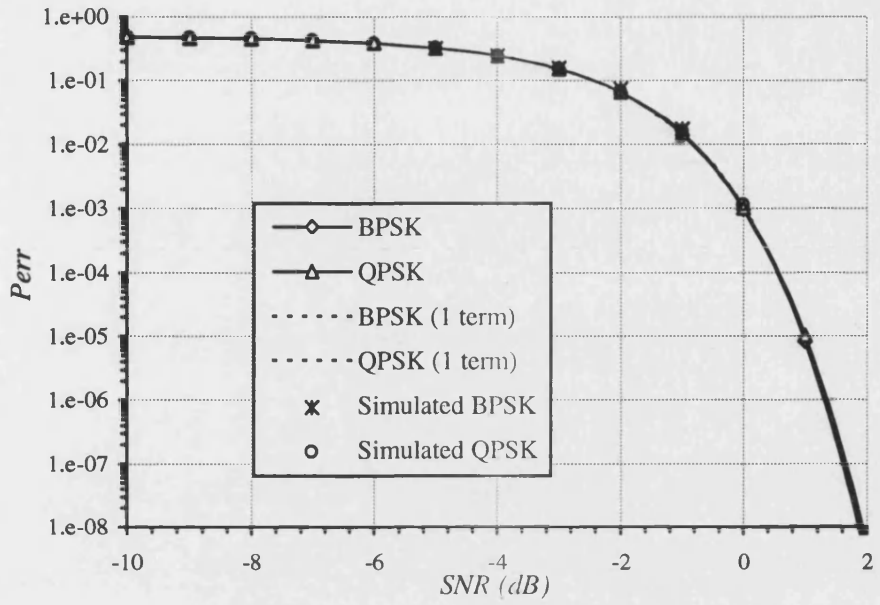


Figure 3.21 : Plot of Misclassification Probability against SNR for the BPSK and QPSK case, L=1024

As the results of a one term classifier are indistinguishable from a multi-term classifier, the one term classifier is proposed for implementation. The structure of a one term BPSK/ QPSK classifier is described in figure 3.22, but can be readily extended to enable a range of PSK signals to be classified, as in all other PSK classifiers.

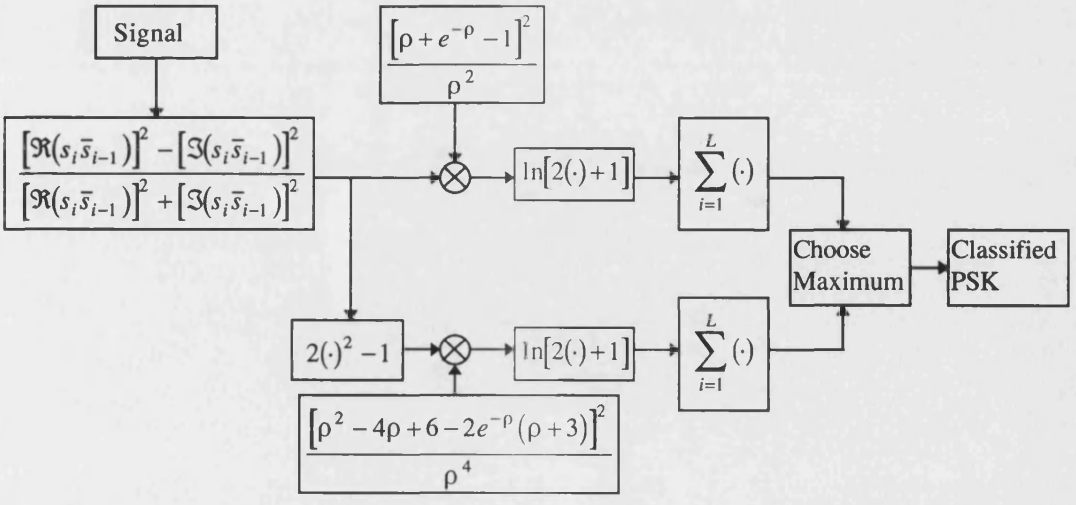


Figure 3.22 : Non-Coherent BPSK/QPSK classifier

3.10.2 DFT of Phase Difference Histogram Classifier

The DFT of phase difference histogram may be achieved in the same way as the DFT of Phase Histogram. The classifier is identical, but the probability of error will be different.

The probability of error is then a modified form of (3.28) to give :

$$p_{err} = \sum_{i=1}^{n} \frac{n!(-1)^{i+1}}{(n-i)!(i+1)!} \exp\left[-\frac{ib_m^4 L}{i+1}\right] \tag{3.106}$$

where there are n+1 classes presented to the classifier. For the BPSK/QPSK classification, the probability of error is given by :

$$p_{err} = \frac{1}{2}\exp\left[-\frac{b_m^4 L}{2}\right] \tag{3.107}$$

The probability of misclassification is given in figure 3.23. It can be seen that a performance penalty is incurred due to the lack of SNR a-priori information.

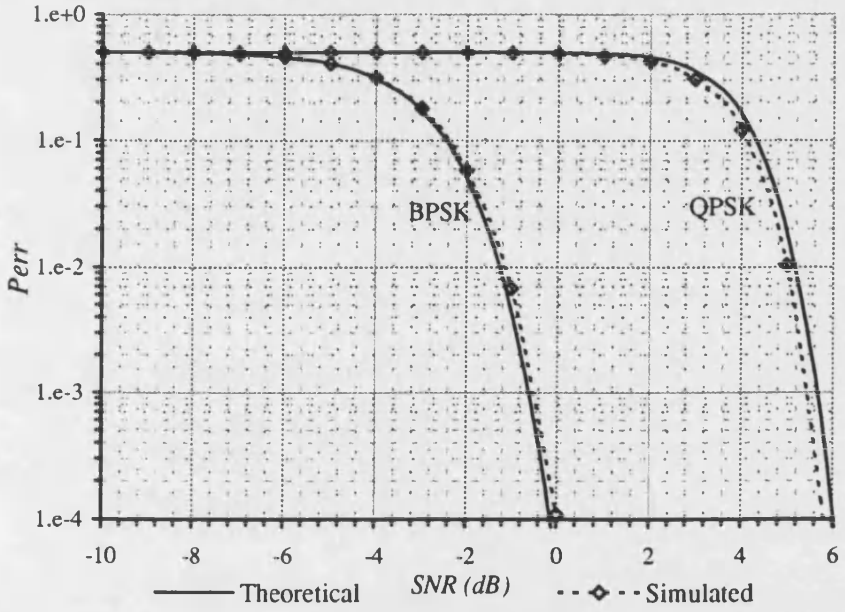


Figure 3.23 : Plots of Misclassification Probability against SNR for the DFT of Phase Difference Classifier, BPSK/QPSK Classified, L=1024

3.10.3 Maximum Likelihood Phase Difference (Non-Modulo 2π) Classifier

In the previous non-coherent techniques, the phase difference was wrapped to modulo 2π. This has the effect of bounding the signal in the interval $(-\pi,\pi]$. The following work develops a classifier which does not apply modulo 2π to the phase difference, and therefore increases the interval to $(-2\pi,2\pi]$.

The statistics of the phase are therefore different, and this work investigates to see if there is any advantage in adopting this approach.

The development starts by modelling the pdf of phase difference for a signal moving from one phase state to another. An average pdf is determined for a given phase shift between samples, which is then extended to M level PSK. From this, a likelihood function is then developed, which is used for classification.

3.10.3.1 Development

The pdf of phase for CW in AGWN has been expressed in terms of a Fourier series (equation 2.7), and the effects of series truncation were observed to improve as the SNR was decreased (Section 2.3.3).

$$f(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{n=1}^{\infty}b_n\cos(n\phi) \qquad \phi\in(-\pi,\pi],\ 0 \text{ otherwise} \tag{3.108}$$

A high SNR approximation for the pdf of phase for a zero phase signal in Gaussian noise may be expressed in terms of the Tikhonov distribution (Section 2.3.4), as :

$$f(\phi)=\frac{\exp[2\rho\cos(\phi)]}{2\pi I_0(2\rho)} \qquad \phi\in(-\pi,\pi],\ 0 \text{ otherwise} \tag{3.109}$$

The non-wrapped phase difference Ω is defined by :

$$\Omega(i)=\phi(i)-\phi(i-1) \tag{3.110}$$

The pdf of Ω may be expressed as the convolution of two pdfs. Assuming that the first phase signal starts at an angle α, and the second at an angle β, the pdf of the phase difference $g(\Omega)$ is given by :

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=f(\phi-\alpha)\otimes f(\phi+\beta) \qquad \Omega\in(-2\pi,2\pi],\ 0 \text{ otherwise} \tag{3.111}$$

which may be expressed analytically as :

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=\int_{-\pi}^{\pi} f(y+\beta)f(\Omega-y-\alpha)dy \tag{3.112}$$

In order to express the convolution without the boundary conditions on ϕ, the following may be used :

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=\int_{\Omega-\pi}^{\pi} f(y+\beta)f(\Omega-y-\alpha)dy \qquad \Omega>0 \tag{3.113}$$

3.10.3.2 MPSK Development

The extension to *M* level PSK requires the analysis of the phase jumps which occur from symbol to symbol. The relative frequency of occurrence of each phase jump is not equiprobable even though each symbol is.

By looking at the phase states of an *M*PSK signal it can be seen that the phase transitions Δ are given by :

$$\Delta \in 0,\pm\frac{2\pi}{M}i \quad i=1\cdots M-1 \tag{3.114}$$

and the relative frequency of occurrence may be deduced by observing the PSK phase states, as in figure 3.24. From this it can be seen that there are *M* possibilities of zero phase change, *M-1* possibilities of $\pm\frac{2\pi}{M}$ phase changes, and so on. The general result is summarised in Table 3.2.

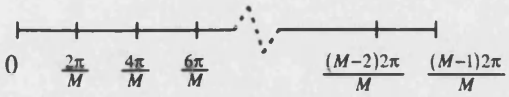


Figure 3.24 : PSK phase states

Phase Transition	0	$\pm\frac{2\pi}{M}$	$\pm 2\frac{2\pi}{M}$	...	$\pm(M-1)\frac{2\pi}{M}$	
Frequency of Occurrence	*M*	*M*-1	*M*-2	...	1	Total $=M^2$

Table 3.2 : Summary of the Relative Frequency of Phase Transitions

The phase transition Δ relates to α and β of equation (3.113) by :

$$\Delta = \alpha - \beta \tag{3.115}$$

The pdf of phase in equation (3.113) may then be written in terms of Δ and β to give :

$$g_{\Delta\beta}(\Omega,\Delta,\beta) = g_{\alpha\beta}(\Omega,\alpha,\beta)_{\alpha=\Delta+\beta} \tag{3.116}$$

As the system is detected in non-carrier synchronous mode, the starting phase β is not known, and will be time varying due to the absence of carrier coherence. The overall pdf is then an average of equation (3.116) with respect to β, where Ω and Δ are constant. This is given by :

$$g_{\Delta}(\Omega,\Delta) = \mathrm{E}\left[g(\Omega,\Delta,\beta),\beta\right] \tag{3.117}$$

The pdf of phase difference $g_M(\Omega)$ for MPSK may then be expressed by :

$$g_M(\Omega) = \frac{1}{M^2}\sum_{i=1-M}^{M-1} g_{\Delta}\left(\Omega,\frac{2\pi i}{M}\right)(M-|i|) \tag{3.118}$$

The general expression for the pdf $g_{\Delta}(\Omega,\Delta)$ in (3.117) is derived in Appendix 3.F, and is given by :

$$g_{\Delta}(\Omega,\Delta,) = \frac{1}{2\pi^2}(\pi - \tfrac{\Omega}{2})\left[1 + \sum_{n=1}^{\infty} b_n^2 \cos[n(\Omega-\Delta)]\right] \qquad \Omega>0 \tag{3.119}$$

This is an infinite series, and a truncated version is used in practice. A three term series may be easily approximated using the Fourier coefficients in (Chapter 2, Appendix 2.B), and this is found to improve as the SNR is decreased (Section 2.3.3) A high SNR approximation is derived in Appendix 3.G, and is given by :

$$g_{\Delta}(\Omega,\Delta) = \frac{1}{2\pi^2\left[I_0(2\rho)\right]^2}(\pi - \tfrac{\Omega}{2})\mathrm{I}_0\left(4\rho\cos\left(\tfrac{\Omega-\Delta}{2}\right)\right) \qquad \Omega>0 \tag{3.120}$$

A suitable switch-over point between the two approximations has been experimentally determined to be approximately 2dB. It should be noted that for negative arguments the pdf is calculated as :

$$g(\Omega<0,\Delta) = g(-\Omega,-\Delta) \tag{3.121}$$

Figure 3.25 shows plots of the pdf of phase difference for BPSK at 2dB SNR for the exact distribution, the high and low SNR approximations to the distribution. This is empirically determined as the changeover point between the high and low SNR algorithms. The approximation can be seen to be good at this worst case.

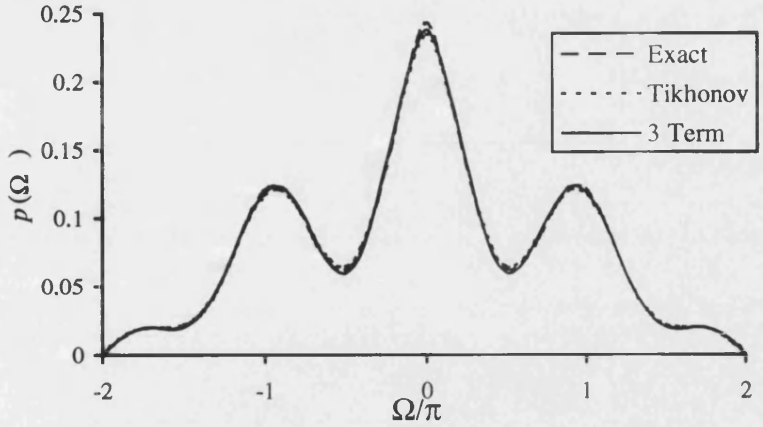


Figure 3.25 : PDF of Phase Difference for BPSK at 2dB SNR

Figure 3.26 shows the pdf of phase difference for QPSK at 6dB SNR for the exact distribution and the high SNR approximation. It is seen that the model is an excellent approximation.

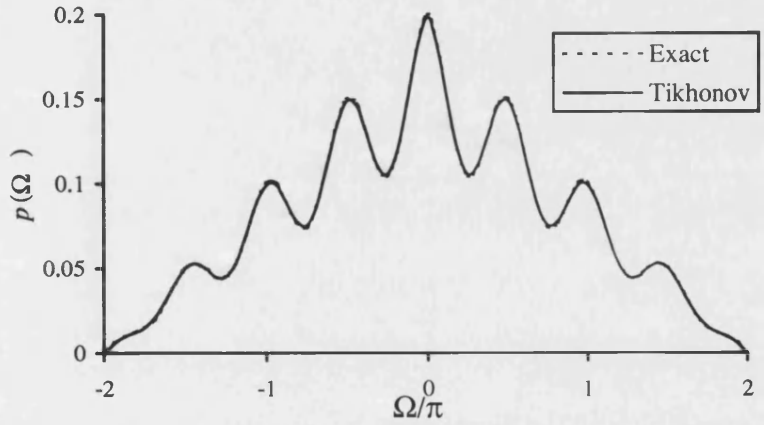


Figure 3.26 : PDF of Phase Difference for QPSK at 6dB SNR

3.10.3.3 Classifier Development

The structure of the classifier is of a maximum likelihood form, where the log-likelihood function for each PSK type is given by the logarithm of the corresponding pdf (Appendix 3.B). Each sample is then passed through the log likelihood function, and summed in turn. The maximum is then reported as the classified PSK type.

The error probability is difficult to determine analytically. In the previous work, the multi-mode nature of the pdf was compensated for by using an aliased version of pdf. This relied

upon symmetry about the modes, which is not present in this case. Therefore the error performance is computed through simulation.

Figure 3.27 shows plots of simulated classification error against SNR when BPSK and QPSK are considered potentially present. From this graph it is seen that the classification performance is similar to that of the classifier based on the phase difference modulo 2π.

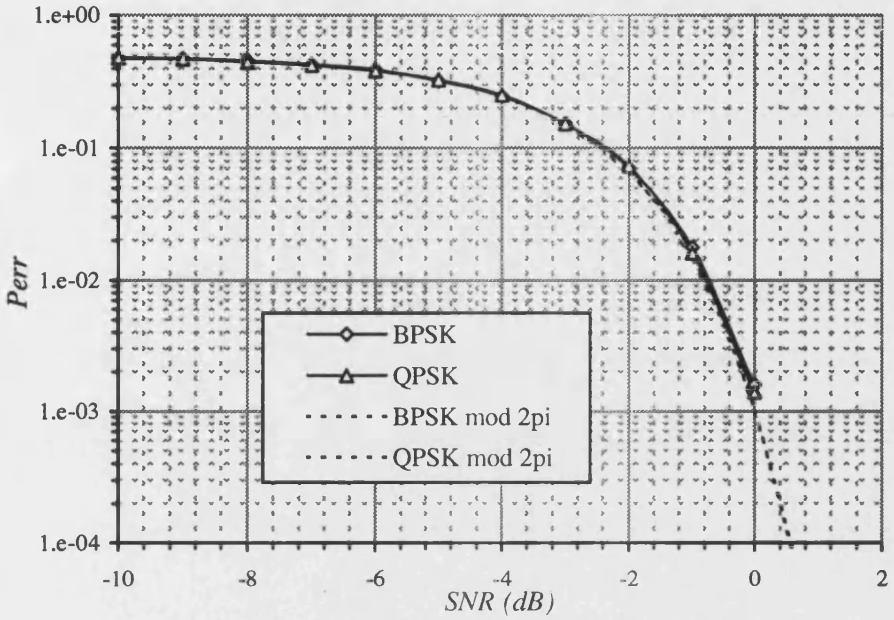


Figure 3.27 : Plots of Misclassification Probability against SNR for the BPSK/QPSK case

3.11 References

[1] A. Polydoros, K.K. Kim, "On The Detection And Classification Of Quadrature Digital Modulations In Broad Band Noise", IEEE Trans. Communications, Vol.38, No.8, pp. 431-436, August 1990.

[2] S.S. Soliman, S.Z. Hsue, "Signal Classification Using Statistical Moments", IEEE trans. Communications, Vol. 40, No. 5, pp. 908-916, May 1992.

[3] Y. Yang and S.S. Soliman, "Statistical Moments Based Classifier For MPSK Signals", Proc. IEEE GLOBECOM, pp. 2.7.1-2.7.5, 1991.

[4] Y. Yang, S.S. Soliman, "An Improved Moment-Based Algorithm for Signal Classification", Signal Processing, Vol. 43, No. 3, 1995, pp. 231-244.

[5] Y. Yang, S.S. Soliman, "Optimum Classifier For M-ary PSK Signals", Proc. IEEE International Conference On Communications, pp. 52.3.1-52.3.5, 1991.

[6] Hwang, C.Y., Polydoros, A., "Advanced Methods for Digital Quadrature and Offset Modulation Classification", IEEE MILCOM '91, pp. 841-845.

[7] C.Y. Huang, A. Polydoros, "Likelihood Methods For MPSK Modulation Classification, IEEE Trans. Communications", Vol. 43, No. 2/3/4, pp. 1493-1504, Feb/March/April 1995.

[8] P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Classification of PSK Signals using the DFT of Phase Histogram", ICASSP-95, Vol. 3, 1995, pp. 1868-1871.

[9] P.C. Sapiano, J.D. Martin, , "Maximum Likelihood PSK Classification using the DFT of Phase Histogram", GLOBECOM '95, vol. 2, pp 1029-1033.

[10] P.C. Sapiano, J.D. Martin, "Maximum Likelihood PSK Classifier", MILCOM '96, to be published.

[11] R. Shiavi, "Introduction To Applied Statistical Analysis", Aksen Associates Incorporated Publishers, Ch. 7, 1991.

[12] I. S. Gradshteyn and I.M. Ryzhik, "Table Of Integrals And Series Products", Academic Press, London, ch. 6.631.4, 1980.

[13] M. Abramowitz, I.A. Stegun, "Handbook of Mathematical Functions, National Bureau of Standards", Washington D.C., Ch. 9,7, 1965.

[14] P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Further Results In The Classification Of PSK Signals Using The Optimum Method", IEE Electronics Letters, Vol. 31, No. 1, pp. 19-20, 1995.

[15] A. Papoulis, "Probability, Random Variables and Stochastic Processes", 2nd ed., McGraw-Hill, 1984, Ch. 7.

[16] A.D. Whalen, "Detection Of Signals In Noise", Academic Press, New York, Ch. 1, 1971.

3.12 Appendix 3.A

Analysis of the Number of Histogram bins in the DFT of Phase Histogram Classifiers

Consider a histogram series $x_1(n)$ of N points in length, and a histogram $x_2(n)$ consisting of the same data but with PN bins. The data from $x_1(n)$ is related to $x_2(n)$ by :

$$x_1(n) = \sum_{m=0}^{P-1} x_2(nP+m) \tag{A.1}$$

The DFT of the two series are :

$$X_1(k) = \sum_{n=0}^{N-1} x_1(n) W^{-kn} \tag{A.2}$$

$$X_2(k) = \sum_{n=0}^{PN-1} x_2(n) W^{-\frac{Kn}{P}} \tag{A.3}$$

where

$$W = \exp j\left(\frac{2\pi}{N}\right) \tag{A.4}$$

It can be shown that :

$$X_2(k) = \sum_{n=0}^{N-1} W^{-kn} \left[\sum_{m=0}^{P-1} x_2(nP+m) W^{-\frac{km}{P}} \right] \tag{A.5}$$

which may be expressed as :

$$X_2(k) = \sum_{n=0}^{N-1} W^{-kn} \left[\sum_{m=0}^{P-1} x_2(nP+m) \exp\left[\frac{-k2\pi m}{NP}\right] \right] \tag{A.6}$$

In this problem $k \leq \frac{N}{4}$, and for $k \ll N$,

$$\sum_{m=0}^{P-1} x_2(nP+m) \exp\left[\frac{-k2\pi m}{NP}\right] \approx \sum_{m=0}^{P-1} x_2(nP+m) = x_1(n) \tag{A.7}$$

Therefore

$$X_2(k) \approx X_1(k) \tag{A.8}$$

This can be seen in figure 3.28, where the lower bins have noticeably similar magnitudes.

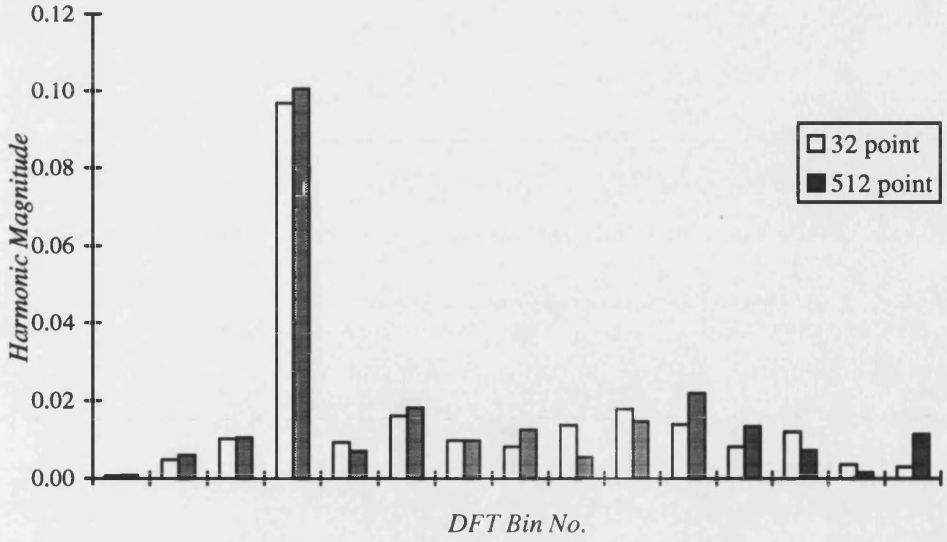


Figure 3.28 : Plots of DFT of Phase Histogram for 32 bin & 512 bin Histograms, QPSK, L=1024, ρ=5dB

3.13 Appendix 3.B

Maximum Likelihood Structure

This section develops the framework for a maximum likelihood classifier, and is used in a number of different classifiers. The content is adapted to the different techniques which are used.

The probability of an observed data sample given the hypothesis that signal type α is transmitted is given by :

$$p(x|\alpha) \tag{B.1}$$

A set of N independent data samples x_i is presented to the classifier, and the joint probability of the data given the hypothesis α is given by :

$$p(x_1, x_2 \ldots x_N|\alpha) = \prod_{i=1}^{N} p(x_i|\alpha) \tag{B.2}$$

The probability of the hypothesis α given the observed data may be determined through Bayes theorem as :

$$p(\alpha|x_1, x_2 \ldots x_N) = p(x_1, x_2 \ldots x_N|\alpha)\frac{p(\alpha)}{p(x_1, x_2 \ldots x_N)} \tag{B.3}$$

The maximum likelihood structure results from the most probable hypothesis i.e. :

$$H_j = \text{MAX}\left[p(\alpha_j|x_1, x_2 \ldots x_N), j\right]_{j\in(1\ldots P)} \tag{B.4}$$

Where P is the number of hypotheses and α_j is the j^{th} hypothesis and H_j is the classified signal. Any constant terms in (B.4) may be removed, as :

$$\text{MAX}\left[p(\alpha_j|x_1, x_2 \ldots x_N), j\right]_{j\in(1\ldots P)} = \text{MAX}\left[K.p(\alpha_j|x_1, x_2 \ldots x_N), j\right]_{j\in(1\ldots P)} \tag{B.5}$$

Assuming that each PSK type is equiprobable, the classified signal may be deduced from (B.5) and (B.4) as :

$$H_j = \text{MAX}\left[p(x_1, x_2 \ldots x_N|\alpha_j), j\right]_{j\in(1\ldots P)} \tag{B.6}$$

Using (B.2), this may be written as :

$$H_j = \mathrm{MAX}\left[\prod_{i=1}^{N} p(x_i|\alpha_j), j\right]_{j\in(1\ldots P)} \tag{B.7}$$

The implementation of (B.7) may run into mathematical overflow when implemented on a digital device. As the logarithm function is a monotonically increasing function, this may be used to reduce the possibility of mathematical overflow as :

$$\mathrm{MAX}\left[\prod_{i=1}^{N} p(x_i|\alpha_j), j\right]_{j\in(1\ldots P)} = \mathrm{MAX}\left[\ln\left[\prod_{i=1}^{N} p(x_i|\alpha_j)\right], j\right]_{j\in(1\ldots P)} \tag{B.8}$$

Rearranging (B.8), classification is based on :

$$H_j = \mathrm{MAX}\left[\sum_{i=1}^{N} \ln\left(p(x_i|\alpha_j)\right), j\right]_{j\in(1\ldots P)} \tag{B.9}$$

The classifier is constructed in a practical system in figure 3.29, where the log likelihood function for each signal class is given by $\lambda_j(x_i)$.

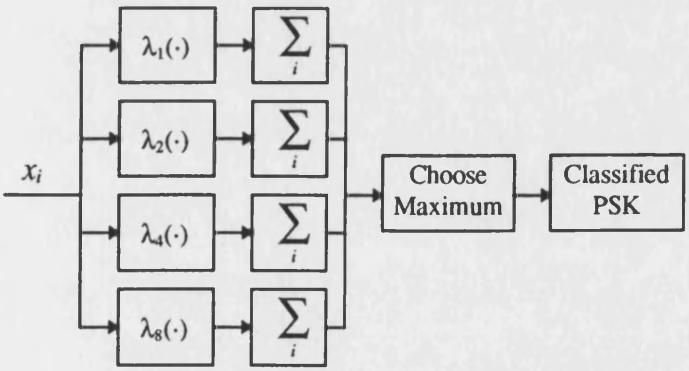


Figure 3.29 : Maximum Likelihood Classifier Structure

3.14 Appendix 3.C

Computational Improvement for the Optimum Phase Classifier

Cubic spline interpolation is an effective technique for evaluating a function when the signal and its derivative are known.

The cubic spline technique evaluates the function f_j and its derivative k_j at discrete points or bins j of spacing h. The value x to be evaluated lies between two bins x_j and x_{j+1} and calculation is performed from the following equation :

$$p_j(x) = a_{j0} + a_{j1}(x - x_j) + a_{j2}(x - x_j)^2 + a_{j3}(x - x_j)^3 \tag{C.1}$$

where :

$$a_{j0} = f_j \tag{C.2}$$

$$a_{j1} = k_j \tag{C.3}$$

$$a_{j2} = \frac{3}{h^2}(f_{j+1} - f_j) - \frac{1}{h}(k_{j+1} + 2k_j) \tag{C.4}$$

$$a_{j3} = \frac{2}{h^3}(f_j - f_{j+1}) + \frac{1}{h^2} \tag{C.5}$$

The pdf of phase for BPSK may be evaluated from (3.50) as :

$$p_2(\phi) = \frac{1}{2\pi} e^{-\rho}\left[\sqrt{\rho\pi}\sin(\phi)e^{\rho\sin^2(\phi)}\mathrm{erf}\left(\sqrt{\rho}\sin(\phi)\right) + 1\right] \tag{C.6}$$

Using [13] the derivative of this function may be evaluated as :

$$p_2'(\phi) = \frac{1}{2\pi} e^{-\rho}\cos(\phi)\left[\sqrt{\rho\pi}\left[1 + 2\rho\sin^2(\phi)\right]e^{\rho\sin^2(\phi)}\mathrm{erf}\left(\sqrt{\rho}\sin(\phi)\right) + 2\rho\sin(\phi)\right] \tag{C.7}$$

For MPSK the pdf is given by :

$$p_M(\phi) = \frac{2}{M}\sum_{i=0}^{\frac{M}{4}-1}\left\{p_2\left[\phi - \frac{\pi(2i+1)}{M}\right] + p_2\left[\phi + \frac{\pi(2i+1)}{M}\right]\right\} \tag{C.8}$$

and the derivative is given by :

$$p_M'(\phi) = \frac{2}{M}\sum_{i=0}^{\frac{M}{4}-1}\left\{p_2'\left[\phi - \frac{\pi(2i+1)}{M}\right] + p_2'\left[\phi + \frac{\pi(2i+1)}{M}\right]\right\} \tag{C.9}$$

It is also possible to incorporate the logarithm function into the interpolation by using :

$$\lambda_M(\phi) = \ln\left[p_M(\phi)\right] \tag{C.10}$$

and :

$$\lambda_M(\phi) = \frac{p'_M(\phi)}{p_M(\phi)} \tag{C.11}$$

Therefore for each trial the log likelihood function and its derivative need only be computed for a number of interpolation points. Each sample is then passed through a very basic set of calculations (C.1) to interpolate the log-likelihood function.

An example of the log likelihood function is given in figure 3.30 for BPSK at 10dB SNR. A 32 bin cubic spline interpolation is applied to this function, and the error is given in figure 3.31. It is seen that the error from this is very low compared to the function itself.

There is some symmetry which may be exploited to reduce the number of coefficients, and occurs about each pdf mode. The function and derivative need only be stored in memory for the range $\left(0, \frac{\pi}{2M}\right)$. However the mapping of this transformation will result in a reduction in computational efficiency.

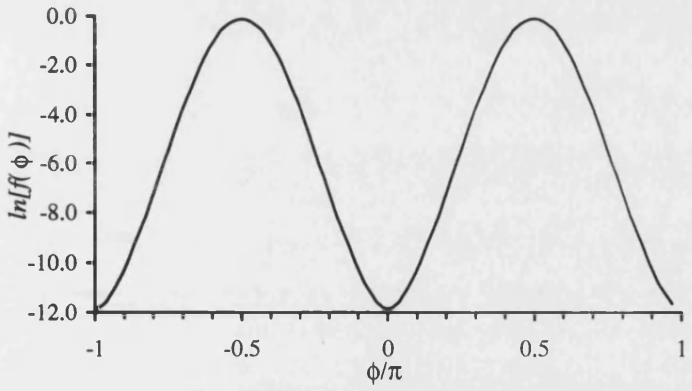


Figure 3.30 : Plot of Log likelihood Function against Phase for BPSK at 10dB SNR

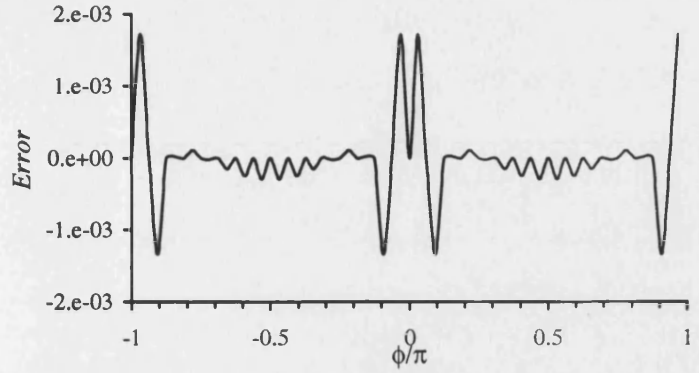


Figure 3.31 : Plot of Interpolation Error against Phase for BPSK at 10dB SNR, 32 Interpolation Bins

3.15 Appendix 3.D

Evaluation of the Classification Performance for the Optimum Phase Classifier

The output of each likelihood function (3.53) is summed over a trial, and the central limit theorem states that this will result in a Gaussian distributed variable when the number of samples is large.

Let α be the PSK type transmitted, and β be PSK type tested. The output of the likelihood function y, is given by :

$$y = \ln\left[p_\beta(\phi)\right] \tag{D.1}$$

Using a transformation of variable [16], the pdf of the output variable is given by :

$$p(y) = p_\alpha\left(\phi|_{\phi=p_\beta^{-1}(e^y)}\right)|J| \tag{D.2}$$

Where J is the Jacobian, defined by :

$$J = \frac{d}{dy}p_\beta^{-1}\left(e^y\right) \tag{D.3}$$

The n^{th} moment m(n,α,β) may then be defined by :

$$\mathrm{m}(n,\alpha,\beta) = \int_{y_{min}}^{y_{max}} y^n p_\alpha\left[p_\beta^{-1}\left(e^y\right)\right]\frac{d}{dy}\left[p_\beta^{-1}\left(e^y\right)\right]dy \tag{D.4}$$

By a change of variable $u = p_\beta^{-1}\left(e^y\right)$:

$$du = \frac{d}{dy}\left[p_\beta^{-1}\left(e^y\right)\right] \tag{D.5}$$

The moments equation is then given by :

$$\mathrm{m}(n,\alpha,\beta) = \int_0^{\frac{\pi}{\beta}} \left(\ln\left[p_\beta(u)\right]\right)^n \tilde{f}_\alpha(u)du \tag{D.6}$$

where $\tilde{f}_\alpha(u)$ is $f_\alpha(u)$ aliased about $\frac{\pi}{\beta}$. This is required for MPSK as the distribution is multi-moded, and therefore the inverse function of (D.4) can only be evaluated within a limited range. The mean and variance of the likelihood function are defined by :

$$\mu_{\beta\alpha} = L\,\mathrm{m}(1,\alpha,\beta) \tag{D.7}$$

$$\sigma^2_{\beta\alpha} = L\left\{\mathrm{m}(2,\alpha,\beta) - \left[\mathrm{m}(2,\alpha,\beta)\right]^2\right\} \tag{D.8}$$

where L is the number of samples. The evaluation of these functions does not lend itself to an analytical solution, therefore the functions are evaluated using numerical integration.

Correct classification occurs when the correct likelihood function is larger than all of the other likelihood functions. The probability of this occurring within the interval $x, x+\delta x$ is given by $k_{\beta\alpha}(x)\delta x$ where :

$$k_{\beta\alpha}(x) = \frac{1}{\sigma_{\alpha\alpha}\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_{\alpha\alpha})^2}{2\sigma^2_{\alpha\alpha}}\right] \prod_{\substack{\beta \\ \beta\neq\alpha}} \Phi\left(\frac{x-\mu_{\beta\alpha}}{\sigma_{\beta\alpha}}\right) \tag{D.9}$$

where :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \tag{D.10}$$

As δx tends towards zero, the overall probability of correct classification is then given by :

$$p_{corr} = \int_{-\infty}^{\infty} k_{\beta\alpha}(x)dx \tag{D.11}$$

3.16 Appendix 3.E

Evaluation of the Likelihood Functions of the Maximum Likelihood IQ Classifier

The likelihood functions of (3.59) may be simplified by the following analysis :

$$\lambda_1(x,y) = \frac{A}{\sigma^2}x \tag{E.1}$$

$$\begin{aligned}\lambda_2(x,y) &= \ln\left[\frac{1}{2}\left\{\exp\left(\frac{A}{\sigma^2}y\right)+\exp\left(-\frac{A}{\sigma^2}y\right)\right\}\right] \\ &= \ln\left[\cosh\left(\frac{A}{\sigma^2}y\right)\right]\end{aligned} \tag{E.2}$$

$$\begin{aligned}\lambda_4(x,y) &= \ln\left[\frac{1}{4}\left\{\exp\left(\frac{A}{\sqrt{2\sigma^2}}[x+y]\right)+\exp\left(\frac{A}{\sqrt{2\sigma^2}}[x-y]\right)+\exp\left(\frac{A}{\sqrt{2\sigma^2}}[-x-y]\right)+\exp\left(\frac{A}{\sqrt{2\sigma^2}}[-x+y]\right)\right\}\right] \\ &= \ln\left[\frac{1}{4}\left\{\left[\exp\left(\frac{A}{\sqrt{2\sigma^2}}x\right)+\exp\left(-\frac{A}{\sqrt{2\sigma^2}}x\right)\right]\left[\exp\left(\frac{A}{\sqrt{2\sigma^2}}y\right)+\exp\left(-\frac{A}{\sqrt{2\sigma^2}}y\right)\right]\right\}\right] \\ &= \ln\left[\cosh\left(\frac{A}{\sqrt{2\sigma^2}}x\right)\cosh\left(\frac{A}{\sqrt{2\sigma^2}}y\right)\right] \\ &= \ln\left[\cosh\left(\frac{A}{\sqrt{2\sigma^2}}x\right)\right]+\ln\left[\cosh\left(\frac{A}{\sqrt{2\sigma^2}}y\right)\right]\end{aligned} \tag{E.3}$$

$$\lambda_8(x,y) = \ln\left[\frac{1}{8}\left\{\begin{matrix} e^{\alpha x}e^{\beta y}+e^{\alpha x}e^{-\beta y}+e^{\beta x}e^{\alpha y}+e^{\beta x}e^{-\alpha y}+ \\ e^{-\alpha x}e^{\beta y}+e^{-\alpha x}e^{-\beta y}+e^{-\beta x}e^{\alpha y}+e^{-\beta x}e^{-\alpha y}\end{matrix}\right\}\right] \tag{E.4}$$

where

$$\alpha = \cos\left(\frac{\pi}{8}\right)\frac{A^2}{\sigma^2} \tag{E.5}$$

$$\beta = \sin\left(\frac{\pi}{8}\right)\frac{A^2}{\sigma^2} \tag{E.6}$$

Therefore :

$$\begin{aligned}\lambda_8(x,y) &= \ln\left[\frac{1}{4}\left\{\left[e^{\alpha x}+e^{-\alpha x}\right]\cosh(\beta y)+\left[e^{\beta x}+e^{-\beta x}\right]\cosh(\alpha y)\right\}\right] \\ &= \ln\left[\frac{1}{2}\left\{\cosh(\alpha x)\cosh(\beta y)+\cosh(\beta x)\cosh(\alpha y)\right\}\right]\end{aligned} \tag{E.7}$$

Note that :

$$\cos\left(\frac{\phi}{2}\right) = \sqrt{\frac{1+\cos(\phi)}{2}} \tag{E.8}$$

$$\sin\left(\frac{\phi}{2}\right) = \sqrt{\frac{1-\cos(\phi)}{2}} \tag{E.9}$$

Therefore :

$$\lambda_8(x,y)=\ln\left[\frac{1}{2}\left\{\cosh\left(a\tfrac{A}{\sigma^2}x\right)\cosh\left(b\tfrac{A}{\sigma^2}y\right)+\cosh\left(b\tfrac{A}{\sigma^2}x\right)\cosh\left(a\tfrac{A}{\sigma^2}y\right)\right\}\right] \qquad \text{(E.10)}$$

where :

$$a=\sqrt{\frac{\sqrt{2}+1}{2\sqrt{2}}} \qquad \text{(E.11)}$$

$$b=\sqrt{\frac{\sqrt{2}-1}{2\sqrt{2}}} \qquad \text{(E.12)}$$

3.17 Appendix 3.F

Low SNR Approximation to the PDF of Phase Difference Non-Modulo 2π

Using (3.108) and (3.113) the auto-convolution may be evaluated as :

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=\frac{1}{\pi^2}\sum_{n=0}^{\infty}\sum_{i=0}^{\infty}b_i b_n\int_{\Omega-\pi}^{\pi}\cos\big(i(\Omega-y-\alpha)\big)\cos\big(n(y+\beta)\big)dy \tag{F.1}$$

where $b_0=½$. This may be written in terms of Δ and β by :

$$g_{\Delta\beta}(\Omega,\Delta,\beta)=\frac{1}{\pi^2}\sum_{n=0}^{\infty}\sum_{i=0}^{\infty}b_i b_n\int_{\Omega-\pi}^{\pi}\cos\big(i(\Omega-y-\Delta-\beta)\big)\cos\big(n(y+\beta)\big)dy \tag{F.2}$$

$$g_{\Delta\beta}(\Omega,\Delta,\beta)=\frac{1}{\pi^2}\sum_{n=0}^{\infty}\sum_{i=0}^{\infty}b_i b_n\int_{\Omega-\pi}^{\pi}\cos\big(i\Omega+(n-i)y-i\Delta+(n-i)\beta\big)+\cos\big(i\Omega-(n+i)y-i\Delta-(n+i)\beta\big)dy \tag{F.3}$$

The average pdf is given by (3.117) as :

$$g_{\Delta}(\Omega,\Delta)=\frac{1}{4\pi}\int_{-2\pi}^{2\pi}g_{\Delta\beta}(\Omega,\Delta,\beta)d\beta \tag{F.4}$$

It can be seen that the cosine terms from (F.3) are zero in all instances where $n\neq i$. The expression may therefore be simplified to :

$$g_{\Delta}(\Omega,\Delta,)=\frac{1}{\pi^2}\int_{\Omega-\pi}^{\pi}dy\sum_{n=0}^{\infty}b_n^2\cos[n(\Omega-\Delta)] \tag{F.5}$$

and is finally given by :

$$g_{\Delta}(\Omega,\Delta,)=\frac{1}{2\pi^2}\big(\pi-\tfrac{\Omega}{2}\big)\left[1+\sum_{n=1}^{\infty}b_n^2\cos[n(\Omega-\Delta)]\right] \tag{F.6}$$

3.18 Appendix 3.G

High SNR Approximation to the PDF of Phase Difference Non-Modulo 2π

Using the Tikhonov approximation (3.109) to the pdf of phase, the pdf of phase difference given angles α and β is deduced from (3.113) as :

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=\frac{1}{4\pi^2[I_0(2\rho)]^2}\int_{\Omega-\pi}^{\pi}\exp[2\rho[\cos(y+\beta)+\cos(\Omega-y-\alpha)]]dy \qquad \text{(G.1)}$$

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=\frac{1}{4\pi^2[I_0(2\rho)]^2}\int_{\Omega-\pi}^{\pi}\exp\left[4\rho\cos\left(\frac{\Omega+\beta-\alpha}{2}\right)\cos\left(y-\frac{\Omega-\alpha-\beta}{2}\right)\right]dy \qquad \text{(G.2)}$$

Using the expansion in [13], this may be re-written as :

$$g_{\alpha\beta}(\Omega,\alpha,\beta)=\frac{1}{4\pi^2[I_0(2\rho)]^2}\int_{\Omega-\pi}^{\pi}I_0\left(4\rho\cos\left(\tfrac{\Omega+\beta-\alpha}{2}\right)\right)+2\sum_{k=1}^{\infty}I_k\left(4\rho\cos\left(\tfrac{\Omega+\beta-\alpha}{2}\right)\right)\cos\left(k\left[\left(y-\tfrac{\Omega-\beta-\alpha}{2}\right)\right]\right)dy \qquad \text{(G.3)}$$

Using the expression in equation (3.116) :

$$g_{\Delta\beta}(\Omega,\Delta,\beta)=\frac{1}{4\pi^2[I_0(2\rho)]^2}\int_{\Omega-\pi}^{\pi}I_0\left(4\rho\cos\left(\tfrac{\Omega-\Delta}{2}\right)\right)+2\sum_{k=1}^{\infty}I_k\left(4\rho\cos\left(\tfrac{\Omega-\Delta}{2}\right)\right)\cos\left(k\left[\left(y-\tfrac{\Omega-\Delta-2\beta}{2}\right)\right]\right)dy \qquad \text{(G.4)}$$

The average pdf is given by (3.117) as :

$$g_{\Delta}(\Omega,\Delta)=\frac{1}{4\pi}\int_{-2\pi}^{2\pi}g_{\Delta\beta}(\Omega,\Delta,\beta)d\beta=\frac{I_0\left(4\rho\cos\left(\frac{\Omega-\Delta}{2}\right)\right)}{4\pi^2[I_0(2\rho)]^2}\int_{\Omega-\pi}^{\pi}dy \qquad \text{(G.5)}$$

$$g_{\Delta}(\Omega,\Delta)=\frac{1}{2\pi^2[I_0(2\rho)]^2}\left(\pi-\tfrac{\Omega}{2}\right)I_0\left(4\rho\cos\left(\tfrac{\Omega-\Delta}{2}\right)\right) \qquad \text{(G.6)}$$

Chapter 4 : Classifier Comparison

4. Classifier Comparison

4.1 Outline

This section attempts to provide a comparison of the different PSK classification techniques. This has previously been made in terms of a-priori information in section 3.2.1, but here a more detailed comparison is given in terms of non-ideal effects where the decision theoretic framework breaks down.

It is not practical to investigate every possible effect or combination of effects, and only a handful have been chosen to illustrate a level of robustness for each technique. These effects include symbol imbalance due to statistical variance, an error in the SNR estimate, a carrier frequency offset and an error in the symbol phase. These are all parameters which may be easily controlled independently.

In order to provide a concise and manageable comparison, the performance of a particular technique is quantified in terms of the SNR at which 1% error classification occurs. In some cases the technique reaches an error floor above 1%. In these cases the performance is not quantified and the SNR at which 1% error probability occurs is graphically placed to infinity.

In the tests 1024 samples have been used throughout, CW, BPSK, QPSK & 8PSK are classified in the coherent techniques and BPSK & QPSK are classified in the non-coherent methods. The Statistical Moments technique uses the 8^{th} statistical moment, which has been found to be suitable for the classification case examined.

	CW	BPSK	QPSK	8PSK	Average
Maximum Likelihood IQ	-19.8	-7.9	0.9	0.9	-6.5
qLLR cs	-19.8	-7.8	1.4	1.4	-6.2
qLLR ns	-19.8	-7	2	2	-5.7
Optimum Phase	-18	-6	1.5	1.5	-5.3
Maximum Likelihood DFT	-17.5	-6	2	2	-4.9
DFT of Phase Histogram	-19	-7	1.5	8	-4.1
8^{th} Statistical Moment	-14	-3.5	5.9	6.1	-1.4

Table 4.1 : SNR (dB) at which 1% Misclassification Probability occurs for the various Synchronous PSK Modulation Classifiers when CW, BPSK, QPSK & 8PSK are transmitted, L=1024

Some of the results presented within this section were presented at IEE RRAS '96 and published in [1]. Table 4.1 summarises the classification performances of the various

synchronous techniques under ideal conditions when CW, BPSK, QPSK and 8PSK are transmitted with 1024 samples.

In this table the classifiers are placed in order of average misclassification performance and it is seen that the Maximum Likelihood IQ classifier has the best performance, as would be expected. The qLLR methods follow this, making the top three performing techniques methods which require estimation of the signal amplitude. The Optimum Phase technique has the best performance out of the phase based methods, closely followed by the Maximum Likelihood DFT of Phase Histogram. Surprisingly the DFT of Phase Histogram method has a better performance than that of the Statistical Moments technique which requires more a-priori information.

These performances are valid under ideal conditions and the following work breaks down some of the assumptions made in the classifier paradigm.

4.2 Finite Effects

In the development of the classification algorithms it was assumed that the number of symbols was equally balanced for each trial. In practice this will not be the case as the symbols will be randomly distributed which results in a relative symbol imbalance. The consequence of this effect is that the pdf on which the classifier is modelled is no longer accurate, and a degradation in performance is expected. Figure 4.1 shows the classifier model for a QPSK pdf, whereas in practice the pdf resembles that of figure 4.2.

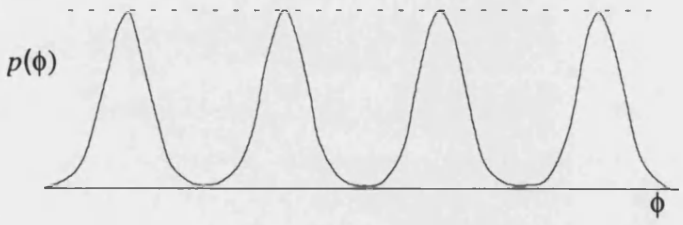


Figure 4.1 : Modelled pdf for QPSK

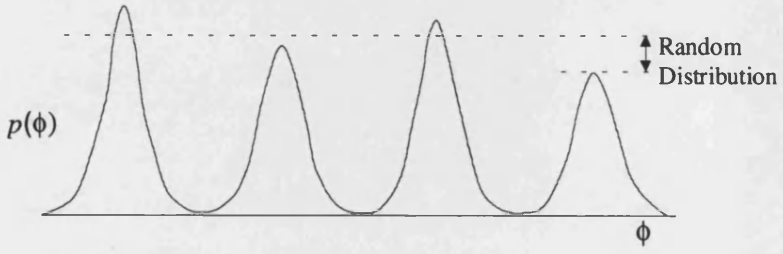


Figure 4.2 : Effect of a Finite Signal Time Frame

Assuming that N symbols are transmitted and the probability of each symbol occurrence is independent and equiprobable, the symbol has a one in M chance of taking a given value. From this definition the number of symbols will be binomially distributed with mean and variance :

$$\mu = \frac{N}{M} \qquad (4.1) \qquad \sigma^2 = N\left(\frac{M-1}{M^2}\right) \qquad (4.2)$$

This has now characterised the distribution of the number of symbols, and the overall distribution of each pdf mode may be determined after normalisation to give :

$$\mu_s = \frac{1}{M} \qquad (4.3) \qquad \sigma_s^2 = \frac{1}{N}\left(\frac{M-1}{M^2}\right) \qquad (4.4)$$

When the number of symbols becomes large, the distribution may be approximated by a normal distribution with mean and variance as above. Table 4.2 shows the 95% confidence intervals for the scaling of the pdf of phase peaks of different PSK signals for different numbers of symbol transmissions.

Symbols	CW	BPSK	QPSK	8PSK
256	0	0.5±0.061	0.25±0.053	0.125±0.041
128	0	0.5±0.0866	0.25±0.075	0.125±0.0573
64	0	0.5±0.1225	0.25±0.1061	0.125±0.081
32	0	0.5±0.1732	0.25±0.15	0.125±0.115

Table 4.2 : 95% Confidence Intervals for PSK pdf Scaling

It is noticed that CW is not subjected to such a variance because there are no symbols present. The performance of the algorithms has been evaluated with 1024 symbols, and the number of symbols is given by the ratio of the total number of samples to the number of samples per symbol. This has been evaluated from 1024 symbols to 32 symbols, and is plotted in figures 4.3-4.6.

The plots for the qLLR cs, Optimum Phase, and the non-coherent phase differencing techniques have been omitted, as the results have found no significant degradation in performance for the range of samples per symbol examined. They are therefore deemed insensitive to this parametric distortion.

From the results in figures 4.3-4.6 it is seen that the Statistical Moments classifier is the most sensitive to this parameter and fails to converge to a 1% classification probability for QPSK and 8PSK at more than four samples per symbol.

The DFT of Phase Histogram method is found to have a gradual degradation for BPSK, QPSK and 8PSK of up to 6dB for 32 samples/ symbol. The qLLR ns method suffers from a significant error with 8PSK, which reaches 4dB degradation at 4 samples per symbol. The Maximum Likelihood DFT method suffers a small degradation on 8PSK, with up to 2dB error for 32 samples per symbol.

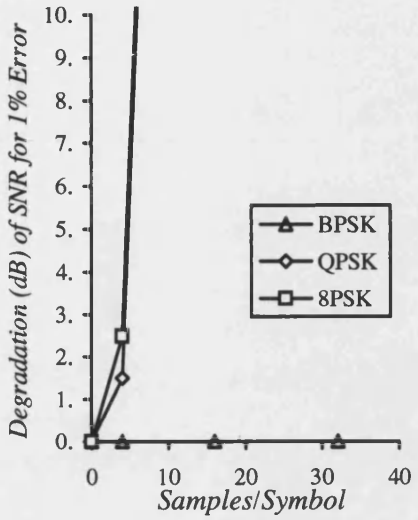


Figure 4.3 : 8th Statistical Moment

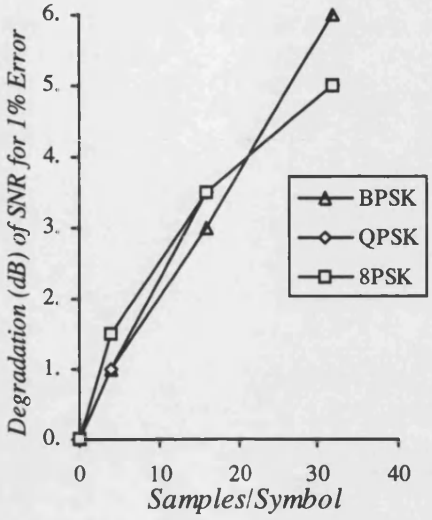


Figure 4.4 : DFT of Phase Histogram

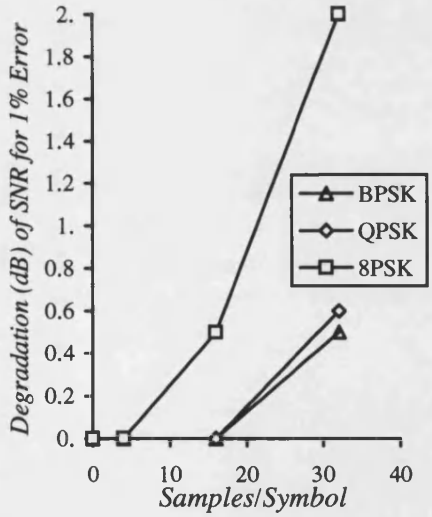


Figure 4.5 : Maximum Likelihood DFT

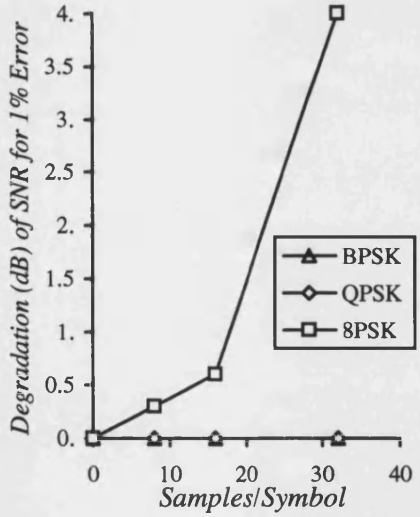


Figure 4.6 : qLLR ns

Plots of SNR for 1% Error Performance against Samples/ Symbol for the 8th Statistical Moments Classifier (Fig. 4.3), DFT of phase Histogram Classifier (Fig. 4.4) Maximum Likelihood DFT Classifier (Fig. 4.5) and qLLR ns Classifier (Fig. 4.6). L=1024

4.3 SNR Offset

Most of the PSK classifiers described require knowledge of the SNR of the incoming signal. This parameter is difficult to estimate with accuracy and it is important to evaluate the sensitivity of the algorithms to this error. The DFT of Phase Histogram method is obviously not affected by this as it makes no use of the signal SNR. This method is therefore used as a bench-mark for comparison.

The qLLR and the Maximum Likelihood IQ methods require both amplitude and SNR knowledge. In order to provide a useful comparison of these three techniques and a more crude comparison of the phase based techniques, the signal is normalised as in [2] to produce a system which is dependent only upon the SNR.

Figures 4.7 to 4.14 show plots of the SNR for 1% classification performance against the ratio of estimated SNR to actual SNR. From these results it is seen that the qLLR cs and ns techniques do not converge when the SNR error is above 1dB for QPSK and 2dB for both BPSK and QPSK. It is also seen that the 8PSK degrades rapidly when the estimate is below the true value.

The Maximum Likelihood DFT method is seen to be the least sensitive of the examined techniques to an SNR offset, with similar performance from the Maximum Likelihood IQ and the Optimum Phase method. The Statistical Moments classifier is seen to have poor performance when the SNR estimate is above the true value, but it is still convergent.

The two non-coherent techniques are seen to converge for the ranges of SNR offset examined, and the symmetrical method is seen to perform better for BPSK when the SNR estimate is above the true value.

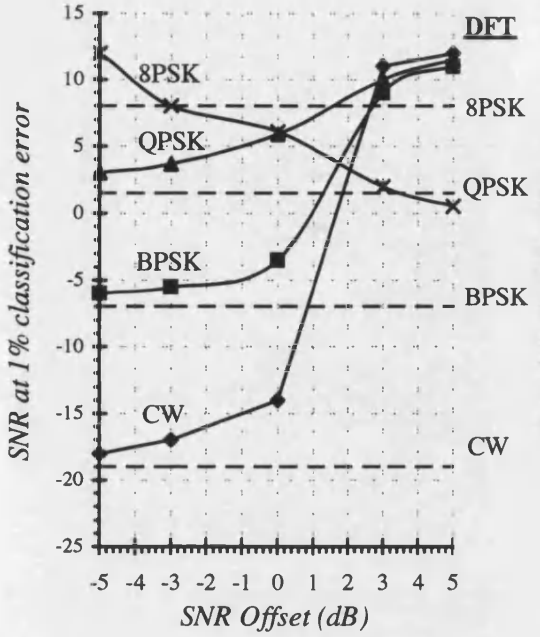


Figure 4.7 : Statistical Moments

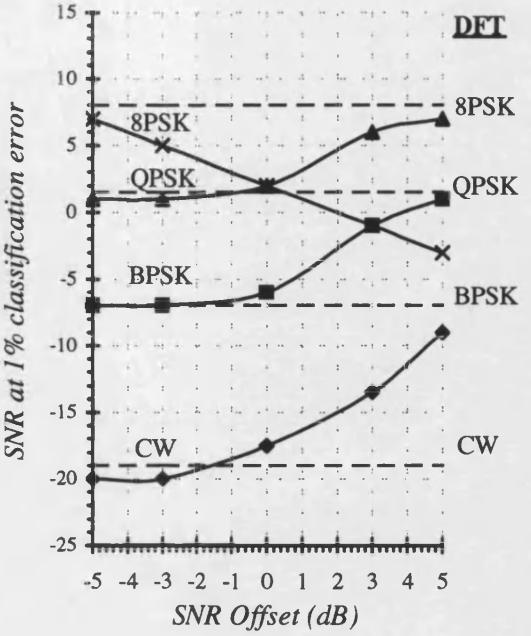


Figure 4.8 : Maximum Likelihood DFT

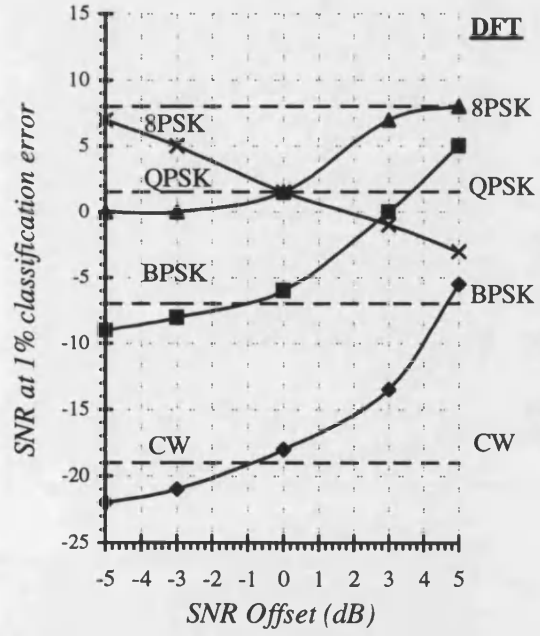


Figure 4.9 : Optimum Phase

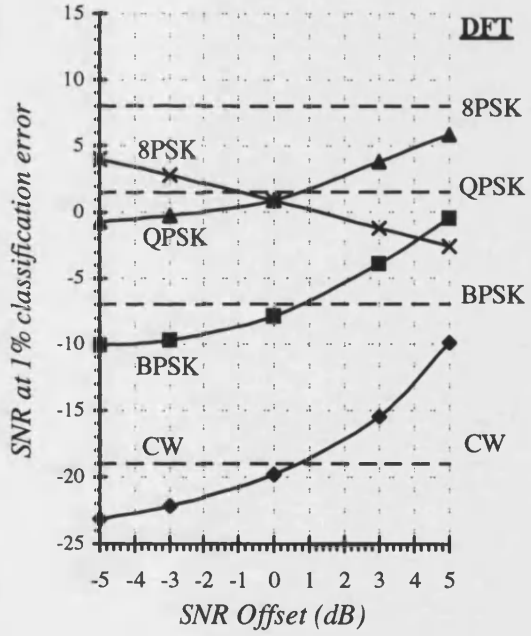


Figure 4.10 : Maximum Likelihood IQ

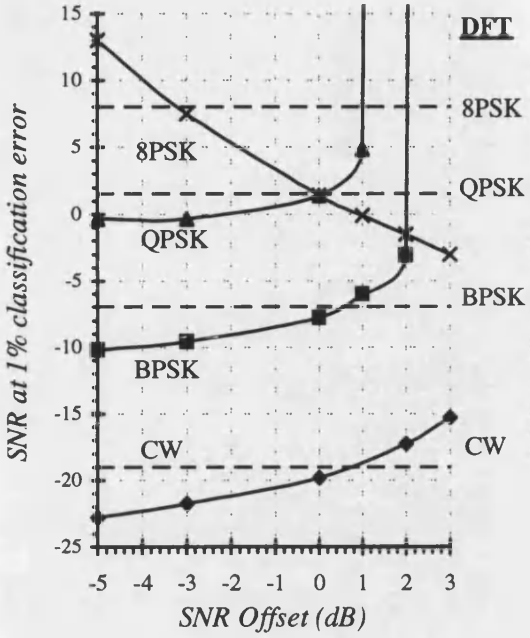


Figure 4.11 : qLLR cs

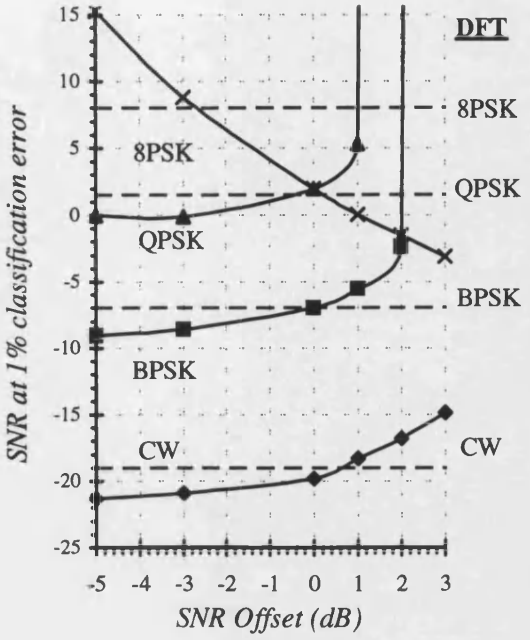


Figure 4.12 : qLLR ns

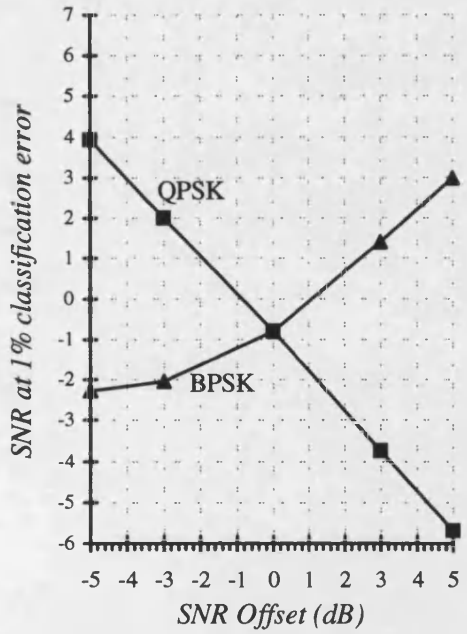


Figure 4.13 : Symmetric Phase Difference

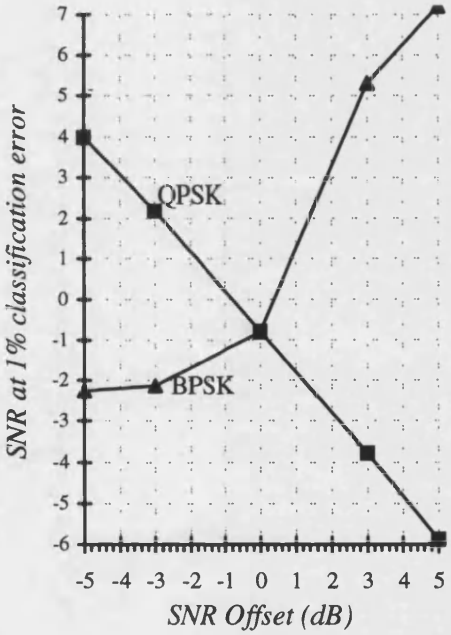


Figure 4.14 : Non-symmetric Phase Difference

Plots of SNR for 1% Error Performance against Ratio of Estimate to True SNR (dB) for the 8th Statistical Moments Classifier (Fig.4.7), Maximum Likelihood DFT classifier (Fig. 4.8), Optimum Phase Classifier (Fig. 4.9), Maximum Likelihood IQ classifier (Fig. 4.10), qLLR cs (Fig. 4.11), qLLR ns (Fig. 4.12), Symmetric Phase Difference (Fig. 4.13) and Non-symmetric Phase Difference (Fig. 4.14)

4.4 Frequency Offset - Coherent Techniques

The synchronisation of the carrier is unlikely to be perfect due to noise and tracking errors, and there will be some variation through the signal analysis time-frame. This will have a significant effect on the synchronous techniques as a frequency offset causes the phase signal to be superimposed upon a linearly increasing phase, resulting in a distorted pdf.

If the frequency offset exhibits a number of cycles during the analysis time frame, the phase will approach a uniform distribution. As the signal observation time is increased, the overall phase shift due to a given frequency error is increased. The measurements which are performed in this section linearly increase the phase offset from zero to a given maximum and henceforth the maximum is termed the 'End Phase Offset'.

Figures 4.15 to 4.21 show plots of the 1% classification performance against End Phase offset for a constant frequency offset, and 1024 points. It is seen that the qLLR cs, Maximum Likelihood IQ and Optimum Phase techniques fail to converge for a maximum phase greater than 0.1π for QPSK, and the Statistical Moments classifier fails to converge for a maximum phase greater than 0.15π. All these techniques require knowledge of the zero phase of the signal, and are expected to perform worse as the zero phase reference is moved with time.

The qLLR ns and the Maximum Likelihood DFT techniques perform well, but fail to converge with QPSK for a maximum phase greater than 0.25π. The DFT of Phase Histogram technique does not perform quite as well, but does perform better than the zero phase synchronised techniques, and fails to converge with 8PSK for a maximum phase offset greater than 0.2π.

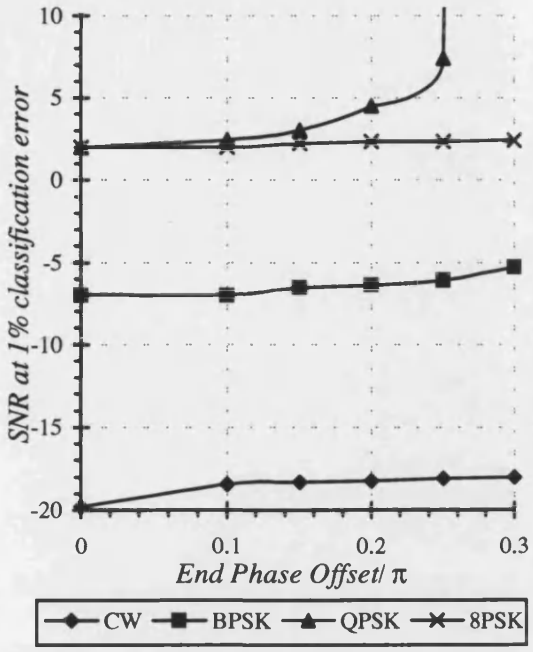


Figure 4.15 : qLLR ns

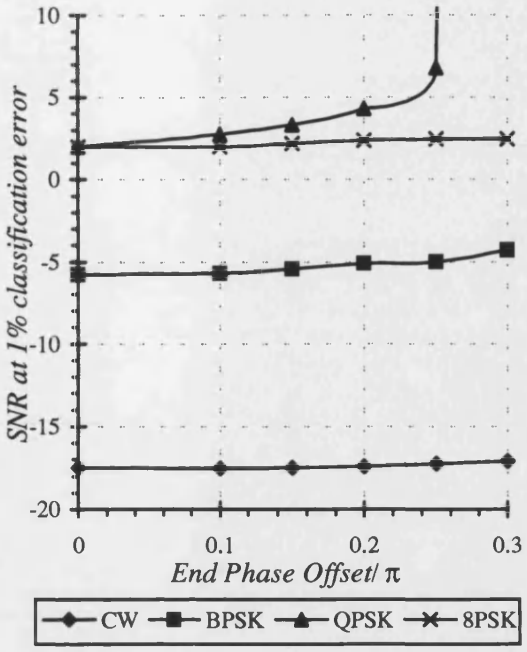


Figure 4.16 : Maximum Likelihood DFT

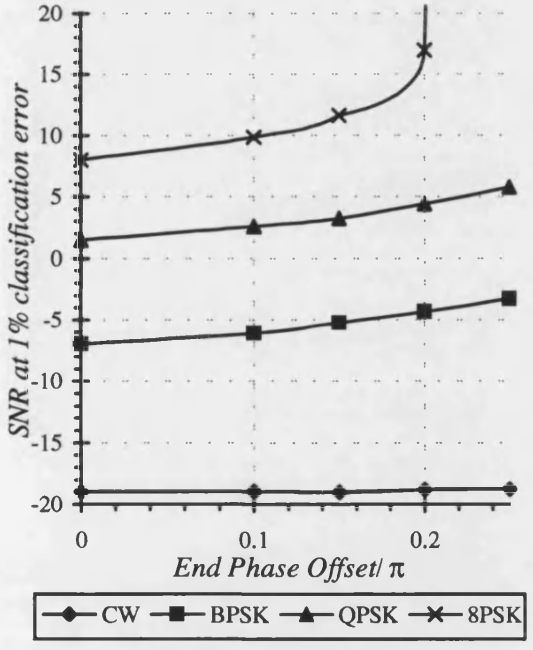


Figure 4.17 : DFT of Phase Histogram

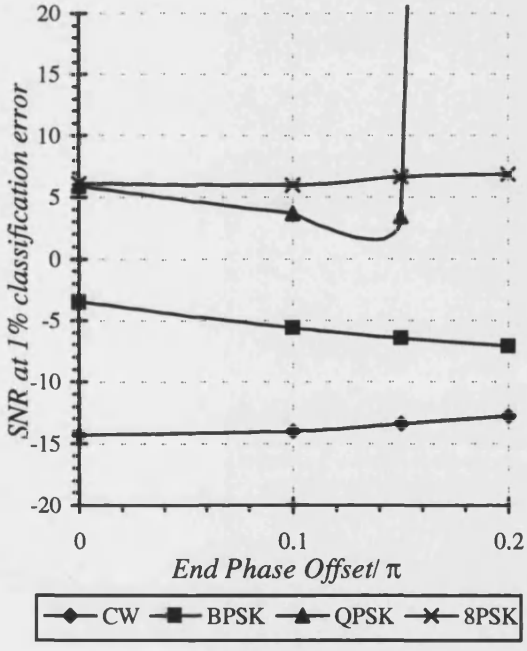


Figure 4.18 : 8^{th} Statistical Moment

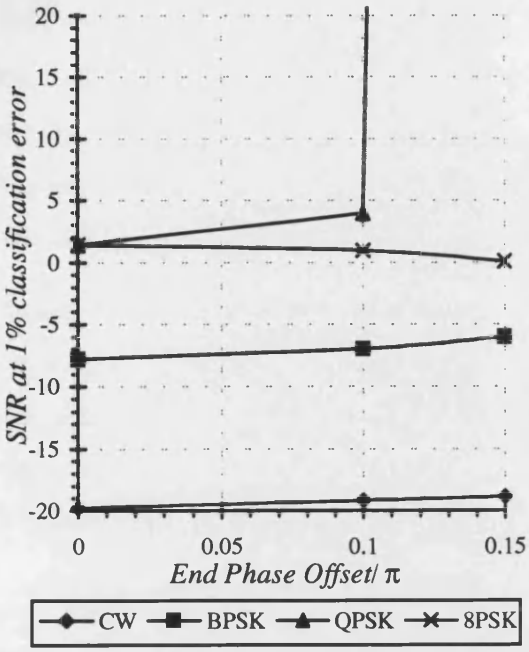


Figure 4.19 : qLLR cs

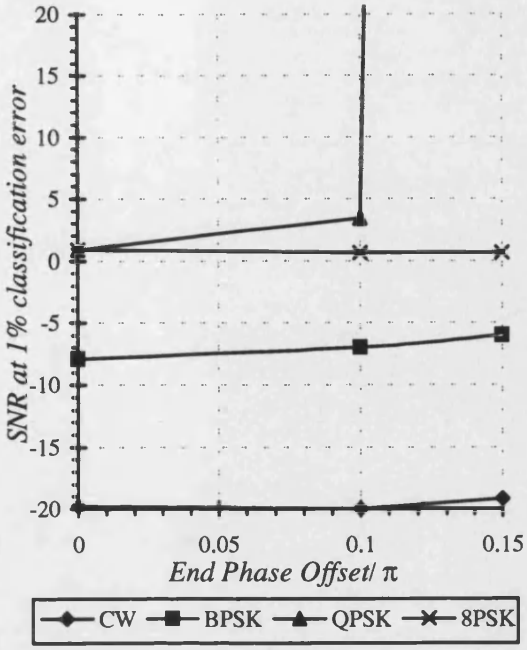


Figure 4.20 : Maximum Likelihood IQ

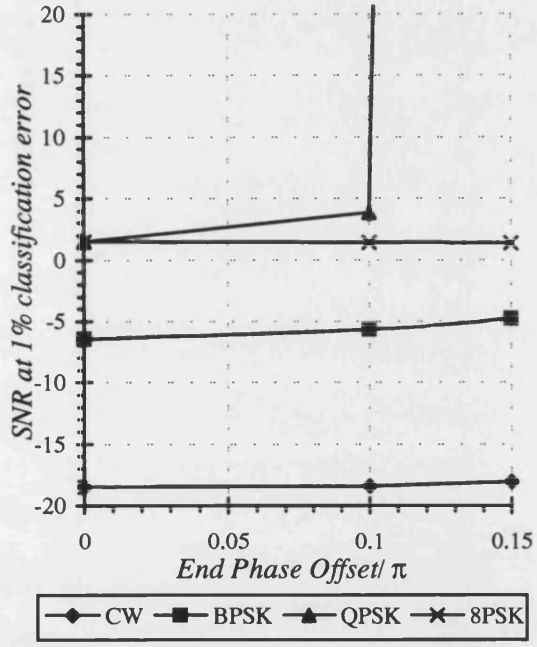


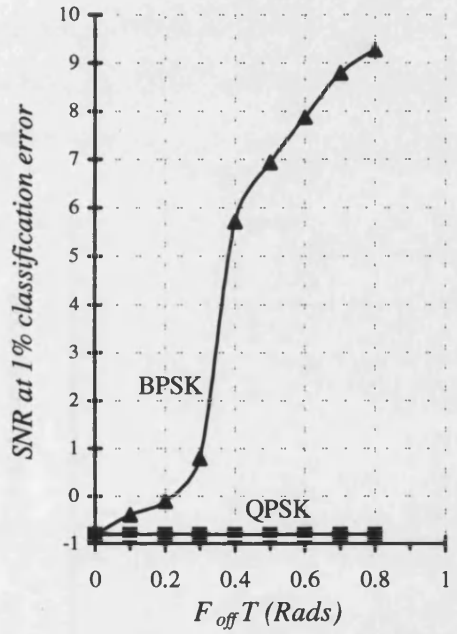
Figure 4.21 : Optimum Phase

Plots of SNR for 1% Error Performance against End Phase Offset for the qLLR ns (Fig. 4.15), Maximum Likelihood DFT Classifier (Fig.4.16), DFT of Phase Histogram (Fig. 4.17), 8th Statistical Moments Classifier (Fig. 4.18), qLLR cs (Fig. 4.19), Maximum Likelihood IQ Classifier (Fig. 4.20), Optimum Phase classifier (Fig. 4.21)

4.5 Frequency Offset - Non-coherent Techniques

The non-coherent techniques are examined under the effect of a frequency offset which is significantly larger than that examined in section 4.4. With the non-coherent classifiers a frequency offset is translated into a constant 'phase' shift in the classifier and hence the pdf does not become distorted.

Figures 4.22 and 4.23 show plots of the SNR for 1% classification performance against frequency offset for the symmetric and non-symmetric phase difference techniques. It is seen that the techniques degrade only on BPSK and they have a similar sensitivity up to approximately 0.3 radians offset. Finally, both techniques become severely degraded above 0.4 radians offset.



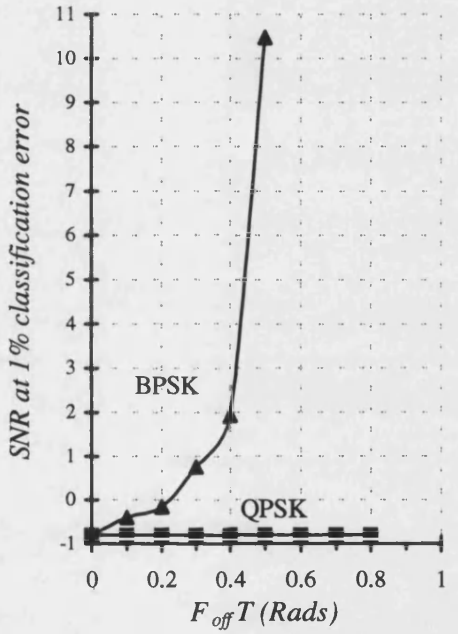


Figure 4.22 : Symmetric Phase Difference *Figure 4.23 : Non-symmetric Phase Difference*

Plots of SNR for 1% Error Performance against Frequency Offset for the non-coherent Symmetric Phase Difference Classifier (Fig. 4.22) and Non-symmetric Phase Difference Classifier (Fig. 4.23)

4.6 Phase Offset

The Statistical Moments, qLLR carrier synchronous, Optimum Phase and Maximum Likelihood IQ techniques all require knowledge of the zero phase of the signal and some of the effects of a changing zero phase reference have been examined in section 4.4.

In a practical system there is likely to be a phase error due to measurement inaccuracies. This error will be a function of the SNR of the signal, and the characteristics will depend upon the type of synchronisation performed. In order to estimate the sensitivity of the different techniques to phase error the error has been fixed across the SNR range, and the performance is again quantified as the SNR at which 1% classification error occurs.

The results have been plotted in figures 4.24 to 4.27 and it is seen that the all the techniques are non convergent above a zero phase error of 12.5°. All of the examined techniques appear to have a similar degree of sensitivity to this parameter, but the statistical moments classifier fails to converge above 10°.

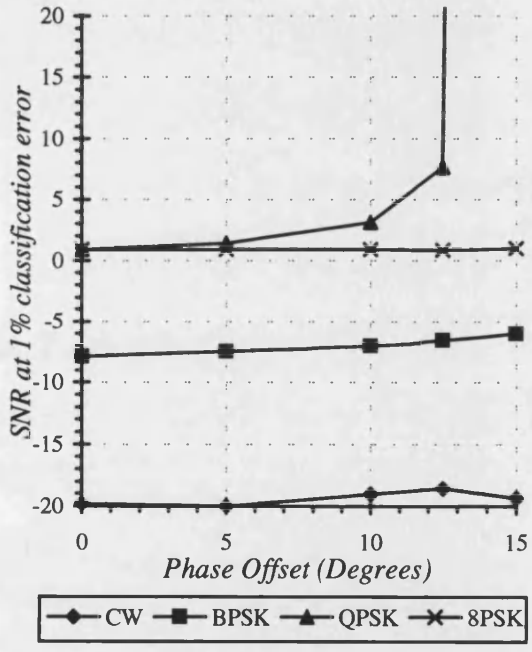


Figure 4.24 : Maximum Likelihood IQ

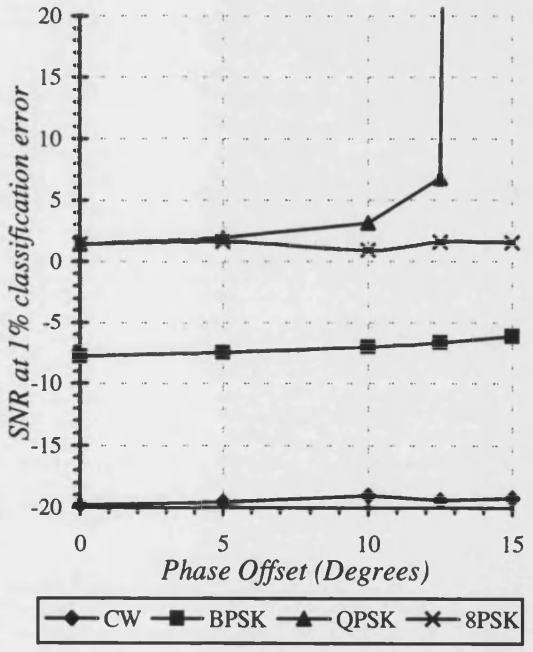


Figure 4.25 : qLLR cs

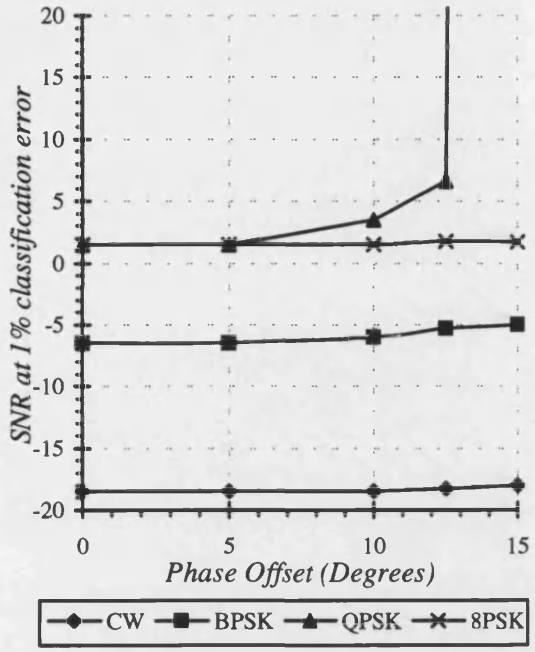


Figure 4.26 : Optimum Phase

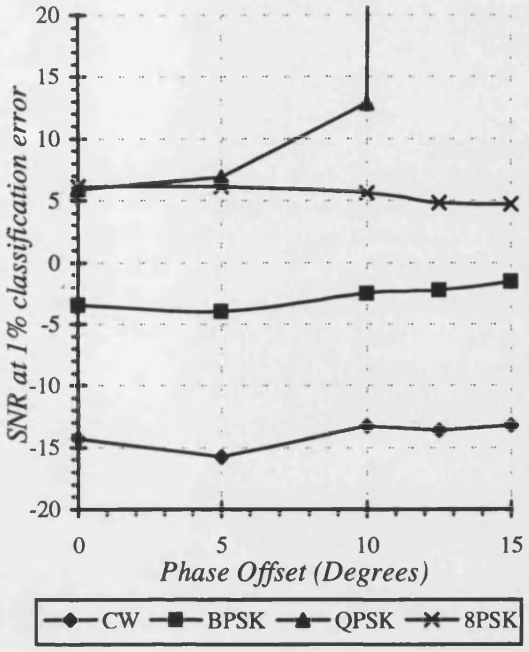


Figure 4.27 : 8^{th} Statistical Moment

Plots of SNR for 1% Error Performance against Static Phase Offset for the Maximum Likelihood IQ Classifier (Fig. 4.24), qLLR cs (Fig. 4.25), Optimum Phase Classifier (Fig. 4.26), 8^{th} Statistical Moments Classifier (Fig. 4.27)

4.7 References

[1] P.C. Sapiano, J.D. Martin, "Identification of PSK Signals", Proc. IEE Radio Receivers and Associated Systems Conference, No. 415, pp. 95-99, September 1995

[2] C.Y. Huang, A. Polydoros, "Likelihood Methods For MPSK Modulation Classification, IEEE Trans. Communications", Vol. 43, No. 2/3/4, pp. 1493-1504, Feb/March/April 1995

Chapter 5 : Conclusions and Future Work

5. Conclusions and Future Work

5.1 Conclusions

This thesis has focused on the classification using decision theoretic techniques, of the number of levels on a PSK signal in additive Gaussian white noise. The foundation of the work has been established through mathematical models which have been developed within the thesis.

The work has produced new synchronous and non-synchronous decision theoretic techniques for PSK classification. The "DFT of Phase Histogram" and "Maximum Likelihood DFT" Classifiers are two new synchronous phase-based PSK classifiers which require different a-priori information to those examined within the published literature. The classification performance of these techniques was found to compare well with other methods and their sensitivity to parametric degradation was less than that for other techniques on a number of aspects.

Three new asynchronous techniques have been developed and their performance degradation compared to the synchronous methods was found to be substantial. However, as these techniques avoid the problem of carrier synchronisation, the associated performance degradation will be acceptable in certain applications.

5.1.1 Models

Chapter 1 has identified that decomposition of the incoming signal into envelope, phase and instantaneous-frequency components, provides three useful signals which may be simply processed to yield quantifiable characteristics of the signal.

In order to obtain these transformations, it was seen in chapter 2 that the signal had to be converted into analytic form. A number of techniques for obtaining a digital analytic signal were investigated in Appendix II. The performance of each was characterised in terms of a rejection ratio of wanted to unwanted components and computational efficiency and it was concluded that the DSP quadrature mixing technique in Appendix I.5.3 achieved the best performance.

It was seen in chapter 2 that the instantaneous-frequency estimate may be achieved through a number of different techniques, each being an approximation to the true instantaneous frequency. A detailed comparison was made in Appendix 2.H of different instantaneous frequency estimators used in the modulation recognition literature and it was concluded that

the finite backward-difference operator was the most suitable estimator for modulation recognition applications.

Using the finite backward-difference operator as the instantaneous-frequency estimator, mathematical models were developed in chapter 2 to characterise the statistics of the envelope, phase and instantaneous-frequency in AGWN. These results are a combination of existing work and new developments, and provide a useful summary for the development of modulation recognition algorithms.

The Fourier series expansion of the pdf of phase for CW and PSK in AGWN has been examined extensively in chapter 2, and iterative numerical methods have been developed for the efficient evaluation of the coefficients in Appendices 2.C and 2.D along with the simplified evaluation of low order Fourier coefficients in Appendix 2.B. These algorithms have been applied to improve the computational efficiency of decision theoretic techniques for PSK classification in sections 3.5, 3.6 and 3.10.

The analysis of the pdf of phase of a PSK signal in section 2.3.3 has developed some simplified properties at low SNR, which have been used extensively in sections 3.4, 3.5 and 3.10 for the decision theoretic classification of PSK signals. New results on the modelling of the pdf of phase difference with and without a modulo 2π operator, have been successfully developed in 2.3.7, where the results were applied to the asynchronous decision theoretic classification of PSK signals in section 3.10.

The results from chapter 2 have provided a useful summary of mathematical models of signal statistics for use in modulation recognition applications. The chapter is useful as a tool in many modulation recognition applications and has provided the foundation for the following chapters, which cover the classification of PSK signals using decision theoretic techniques.

5.1.2 PSK Classification

There have been a number of aspects of PSK classification addressed within the thesis. Some work has been presented which streamlines existing techniques (Section 3.6, 3.8.3 and Appendix 3.C) and provides new areas of performance analysis (Section 3.8.2 and Appendix 3.D) , but the most important areas of work have been the generation of new decision theoretic techniques.

5.1.2.1 Carrier frequency Synchronous Techniques

The DFT of phase histogram and maximum likelihood DFT of phase histogram methods (Sections 3.4 and 3.5) are two new carrier synchronous methods for classifying the number of levels on a PSK signal. Both techniques are insensitive to a reference phase offset, which will be important in a practical system as most carrier synchronisation methods produce a phase reference ambiguity.

The classification performance of the DFT of phase histogram has been successfully modelled (Sections 3.4.2-3.4.5), the maximum likelihood DFT of phase histogram has been simulated, and the classification performance of both techniques has been found to be good in comparison with other methods.

The global maximum likelihood technique (Maximum Likelihood IQ Classifier) has been developed in section 3.8 and bases classification on the in-phase and quadrature signals. The a-priori information required for the different techniques is compared in table 3.1, where the maximum likelihood IQ classifier was shown to require knowledge of the carrier phase reference, SNR and amplitude information. In contrast, the two DFT methods do not require amplitude and carrier phase reference knowledge, and the DFT of phase histogram method does not require SNR knowledge.

The classification performance of the various techniques has been compared under ideal conditions in table 4.1, and it has been found that the average performance loss for a CW-8PSK classification problem is 1.6dB for the maximum likelihood DFT of phase histogram and 2.4dB for the DFT of phase histogram when compared to the Maximum Likelihood IQ classifier.

The qLLR ns method was the only method found in the literature which does not require a carrier phase reference. However, it was found to require more a-priori knowledge than the DFT methods, requiring knowledge of the amplitude and SNR of the signal. From Table 4.1 the maximum likelihood DFT method was found to perform 0.8dB worse than the qLLR technique and the DFT of phase histogram method performed 1.6dB worse.

Some issues relating to the classifier sensitivity were examined in chapter 4. The first issue examined was the performance degradation due to statistical imbalance of the number of different symbols (Section 4.2), which violated the classifier assumption that an equal number of symbols are present in the analysis time frame. It was found that some classifiers were invariant to this imbalance. The statistical moments classifier was found to be particularly

sensitive to this effect (figure 4.3), more gradual degradations were exhibited in the qLLR ns and the DFT of phase histogram (figures 4.4 and 4.6), and the maximum likelihood DFT was only slightly affected (figure 4.5).

The effects of an error in the SNR estimate were examined in section 4.3, and the DFT of phase histogram method was found to be invariant to this parameter. Figures 4.7-4.12 show that when the SNR estimate was above the true value CW-QPSK degraded and 8PSK improved, but the opposite was true when the SNR estimate was below the true value. The qLLR techniques were discovered to be particularly sensitive to this offset (figures 4.11 and 4.12), especially when the SNR estimate was above the true value. The statistical moments classifier was also found to be sensitive to the offset (figure 4.7), but the other classifiers had less sensitive, similar characteristics.

The effect of a linear phase offset was examined in section 4.4. All the techniques which did not require a phase reference, failed to converge for QPSK when the end phase offset was 0.2π-0.25π radians (figures 4.15-4.17). The other techniques failed to converge above 0.1π radians (figures 4.18-4.21)

From the comparative sensitivity measurements of chapter 4, it was found that the maximum likelihood DFT classifier was on balance one of the least sensitive to parametric error. The DFT of phase histogram technique was found to perform well compared with the other techniques.

Based on their performance, required a-priori information, sensitivity and computational efficiency, the DFT of phase histogram and maximum likelihood DFT of phase histogram techniques represent two useful techniques in the development of carrier frequency synchronous decision theoretic PSK classification techniques.

5.1.2.2 Asynchronous Techniques

Three new carrier frequency non-synchronous methods of PSK classification have been developed in section 3.10 using decision theoretic techniques.

Two maximum likelihood structures were developed (Sections 3.10.1 and 3.10.3), one using the phase difference modulo-2π and the other using the phase difference which is not subjected to modulo-2π arithmetic (Section 3.10.3). The pdf structures of the two methods were found to be substantially different and the phase difference modulo-2π technique was found to be significantly simpler in structure.

The performance of the techniques was compared in figure 3.27 and it was found that there was little perceivable difference in the classification performances. The penalty incurred through the use of an asynchronous compared to a frequency synchronous technique was found to be around 7dB when compared to that of the maximum likelihood IQ method in figure 3.14.

The sensitivity of the asynchronous techniques to the parametric distortion applied in chapter 4 was found to be similar for both techniques, and they were both insensitive to symbol imbalance. This sensitivity to an SNR offset was found to be similar to the better synchronous techniques and their sensitivity to a frequency offset is characterised in figures 4.22 and 4.23.

The maximum likelihood phase difference modulo 2π was found to be the preferred technique, mainly due to its considerably simplified structure (figure 3.22).

The DFT of phase histogram structure was extended to the DFT of phase difference histogram classifier in section 3.10.2 and the classification performance was accurately modelled (figure 3.23). This classifier does not require knowledge of the SNR and is intolerant to small frequency offsets, but has a 3dB average penalty associated with it for a BPSK/QPSK classification case.

As the receiver structure for the asynchronous techniques does not require carrier synchronisation, the complexity is reduced considerably. The discussion in Appendix I has highlighted the difficulty of achieving carrier synchronisation when the number of PSK levels are unknown. It has also pointed out that the problems associated with blind symbol synchronisation are not as great, therefore the asynchronous methods provide a useful method for PSK classification despite the associated performance penalty.

5.2 Future work

There is much work still required in decision theoretic PSK classification to take it from a theoretical technique to a practical one and some direction is provided in Appendix I for a number of these issues.

Using carrier frequency synchronous methods of classification, the ability to generate a carrier reference has been assumed. Some of the published literature [1-8] has attempted to provide techniques for carrier frequency estimation, but none provide sufficient accuracy for PSK type applications under noisy conditions. Other methods for blind synchronisation have been touched upon in this thesis in Appendix I, but further work is required to produce a robust method.

In both synchronous and asynchronous carrier frequency techniques it is desirable to have knowledge of the symbol rate and timing information. A blind timing recovery system is conceptually more simple to achieve than carrier recovery and has been discussed in brief in Appendix I of this thesis, but further work is clearly required in this area. Some work is also presented within the published literature for timing recovery [6-9] with the most valuable contribution from [9].

Once a set of carrier and symbol synchronisation techniques are developed for PSK classification, it will be important to evaluate the overall system performance with the effects of synchronisation errors, which will give a clearer idea of the system performance.

In a broad modulation recognition problem it has been identified that the generic split between analogue and the different digital methods is important, and research into such techniques is required. This then enables the decision theoretic methods to be evaluated independently of other modulation schemes outside the scope of the classifier. Some techniques for achieving this will most likely come from analysis of the cyclostationary characteristics of the digitally modulated signals which occur in the symbol timing information.

PSK classification has now been researched in some depth, and still requires further development. However, there is a lack of decision theoretic classification techniques for other forms of digitally modulated signals. Effort may then be spent in unifying the techniques such that e.g. FSK and PSK decision theoretic techniques may be combined.

Methods for estimating the SNR of the signal require some development. The methods of deriving the SNR from the envelope statistics (Appendix I) require an analysis under non-ideal conditions where the noise is not white.

The PSK classification schemes have been evaluated under the effects of AGWN and under a limited number of conditions within this thesis. It would be useful in future work to evaluate the effects under non-Gaussian noise, distortion, co-channel interference and multi-path.

In order to progress the research of analogue modulation recognition it is envisaged that speech characteristics are processed and analysed in more depth so that the pattern classifier can utilise the types of characteristic which a human is capable of discriminating.

The key to such an area may be held within the research into automatic speaker or language recognition, and could be a challenging problem to tackle. Some initial work into the area has been performed in [10].

5.3 References

[1] L. Vergara Dominguez, J.M. Páez Borrallo, J. Portillo Garcia, "A General Approach to the Automatic Classification of Radiocommunication signals", Signal Processing, Vol. 22, 1991, pp. 239-250.

[2] K.R. Farrell, R.J. Mammone, "Modulation Classification Using a Neural Tree Network", MILCOM '93, pp.1028-1032.

[3] R.J. Inkol, R.H. Saper, "A New Algorithm for Signal Classification", Proc. NAECON '93, pp. 320-326.

[4] J.E. Whelchel, D.L. McNeill, R.D. Hughes, M.M. Loos, "Signal Understanding : an Artificial Intelligence Approach to Modulation Classification", Proc. IEEE International Workshop on Tools for A.I. 1989, pp. 231-236.

[5] P.A.J. Nagy, "A Modulation Classifier for Multi Channel Systems and Multi Transmitter Situations", IEEE MILCOM '94, Vol. 3, pp. 816-820.

[6] R.J. Mammone, R.J. Rothaker, C.I. Podilchuk, S. Davidovici, D.L. Schilling, "Estimation of Carrier Frequency, Modulation Type and Bit Rate of a Modulation Signal", IEEE ICC '87, pp. 28.4.1- 28.4.7.

[7] S.Z. Hsue, S.S. Soliman, "Automatic Modulation Classification Using Zero Crossing", IEE Proceedings, Vo. 137, Pt. F, No. 6, Dec. 1990, pp. 459-464.

[8] K. Assaleh, K.R. Farrell, R.J. Mammone, "A New Method of Modulation Classification for Digitally Modulated Signals", MILCOM '92, pp. 30.5.1-30.5.5.

[9] F.F. Liedtke, "Computer Simulation of an Automatic Classification Procedure for Digitally Modulated Communication Signals with Unknown Parameters", Signal Processing, Vol. 6, No. 4, August 1984, pp. 311-323.

[10] J.D. Hoyt, H. Wechsler, "Detection of Human Speech in Structured Noise", Proc. ICASSP '94, Vol. 2, pp. 237-240.

Chapter 6 : Acknowledgements

6. Acknowledgements

The Author would like to express thanks to the following people for their assistance within this work :

I would like to thank R.J. Holbeche who co-supervised this PhD. for almost two years until he unfortunately had to retire early on medical grounds.

I would like to express particular thanks to J.D. Martin who has provided excellent support in my work.

I would also like to thank my family for their long term support and encouragement throughout my academic life.

Finally I would like to thank my new wife Susan, for her support throughout the PhD and for her assistance in the proof reading of this thesis.

This work was supported by Vodafone plc. and EPSRC.

Appendix I : Techniques for Synchronisation and Parametric Estimation

I. Techniques for Synchronisation and Parametric Estimation

I.1 Outline

This section provides a basic insight into the way in which decision theoretic PSK classification schemes may fit into a practical system.

It has been found that there is a lack of detail in the published literature addressing the problems of carrier synchronisation, symbol timing recovery and parametric estimation for PSK classification techniques. The work provided in this chapter provides merely a qualitative discussion of synchronisation and parametric estimation techniques, and helps to provide some direction for future work into these issues for decision theoretic PSK classification techniques.

The development of a modulation blind carrier synchronisation algorithm is by no means a trivial exercise, and there is no known work present within the published literature addressing such a problem.

A modulation blind PSK carrier reference is shown in section I.2.1., where the analytic PSK signal is taken to a power equivalent to the maximum number of levels potentially present on the incoming signal. This then provides a discrete frequency component at a multiple of the carrier frequency which may be used for synchronisation. However, the performance of such a synchronisation technique is poor for the PSK classes less than the maximum PSK scheme. This example demonstrates that a modulation blind PSK synchronisation algorithm is achievable, but research is required in order to improve the performance of such a technique.

Other techniques for implementing carrier frequency synchronous modulation recognition methods without a modulation blind synchroniser are discussed. One is detailed in section I.2.2 where power law synchronisation is applied for each PSK type which is assumed potentially present. The estimated carrier is then removed and the modulation recognition algorithm is applied to each case. The modulation recognition algorithm then reflects a degree of confidence for a particular PSK type, and the most likely PSK type is chosen.

It should be noted that these methods of carrier synchronisation introduce a phase ambiguity into the recovered carrier reference. The effect of this has implications on the implementation

of the qLLR cs, maximum likelihood IQ, statistical moments and optimum phase classifiers, where this phase ambiguity is unacceptable (Section 4.6).

Section I.3 shows that modulation blind symbol timing recovery is a great deal more straight forward than carrier recovery and three different techniques to generate a reference signal are discussed. The structure of each method is simple and remains the same for the different ranges of PSK signal which may be present.

Finally, section I.4 details some methods for SNR and amplitude estimation. One of the methods is based on the receiver noise floor and the other is based upon the envelope statistics of the detected signal.

I.2 Carrier Synchronisation

The synchronous techniques for PSK classification require accurate knowledge of the carrier frequency of the signal. The accuracy of this estimate is required such that the overall phase offset at the end of a signal time frame is significantly less than 2π radians. i.e.

$$L\omega_{off}T_s << 2\pi \tag{I.1}$$

where ω_{off} is the offset angular frequency, T_s is the sampling interval and L is the number of samples. Section 4.6 shows the sensitivity of the various PSK classification techniques to a carrier frequency offset and it is seen that for a 1024 sample time frame, the maximum tolerable end of phase offset is 0.25π for the qLLR ns and maximum likelihood DFT techniques.

The accurate estimation of the carrier frequency of a signal is an extremely difficult process when the modulation type is unknown, and this is the most significant drawback of the carrier frequency synchronous methods for PSK classification.

Techniques for the generation of a carrier reference have been extensively studied within communication systems theory and these generally rely upon the generation of a spectral component resulting from a transformation of the signal. A common method is to place the signal through a power law, which yields a discrete spectral component at multiples of the carrier frequency. This is then followed by a phase tracking device such as a phase locked loop to track the carrier frequency.

In modulation recognition, such a phase locking device is not always suitable, as the signal observation is of a limited duration and the device does not provide a confidence metric for the locking process. An alternative technique is to use spectral methods to detect the discrete

component. One such method is the DFT, which has been found to be the optimum method for detecting a single sinusoid in AGWN [1]. Due to the simplicity of the DFT, this is an obvious starting point for the spectral estimation.

For a PSK signal, the n^{th} law will rotate each constellation point by n times the phase. It is found that with a noise free rectangular signal the constellation of MPSK wraps to a single point when $n \geq M$. This produces a discrete spectral line at n times the carrier offset frequency. The power law applied in the communication systems, often uses an n^{th} law device on the 'real' signal. This does however have some undesirable effects due to unwanted components when $n>2$. These components are generated on a noise free signal and are demonstrated from (I.2).

$$[A\cos(\phi)]^n = \underbrace{\frac{A^n}{2^{\frac{n}{2}-1}}\cos(2n\phi)}_{\text{Wanted}} + \underbrace{\frac{1}{2^{\frac{n}{2}}}\left\{\sum_{k=1}^{\frac{n}{2}-1} 2\binom{n}{k}\cos\left[2\left(\tfrac{n}{2}-k\right)\phi\right] + \binom{n}{k}\right\}}_{\text{Unwanted}} \tag{I.2}$$

A better approach is to apply the n^{th} law device to the analytic signal, which does not suffer from the interference effects of (I.2). This should be achieved within the DSP, as it is difficult to implement in an analogue structure, which is the reason that the method of (I.2) has been extensively used in the past.

$$[A\cos(\phi) + jA\sin(\phi)]^n = A^n\cos(n\phi) + jA^n\sin(n\phi) = A^n\exp[jn\phi] \tag{I.3}$$

The transformation in (I.3) has been applied in [2]. Appendix I.A shows that the detection of the spectral line is worse than the performance of the decision theoretic classifiers. Therefore the decision theoretic classifiers may be used as an additional test of a PSK type being present.

Two methods linking this form of carrier recovery and decision theoretic PSK classification are discussed in the following sections.

I.2.1 Modulation Blind PSK Carrier Recovery

The method of carrier recovery shown in figure I.1 establishes a carrier component based upon the analytic signal transformed with a power law equivalent to the maximum PSK number e.g. if 8PSK is the maximum PSK type assumed present, then an 8^{th} power is applied.

The discrete component is then divided by the power order and is used as the carrier reference, from which the phase may be detected. This is then followed by the PSK classification algorithm.

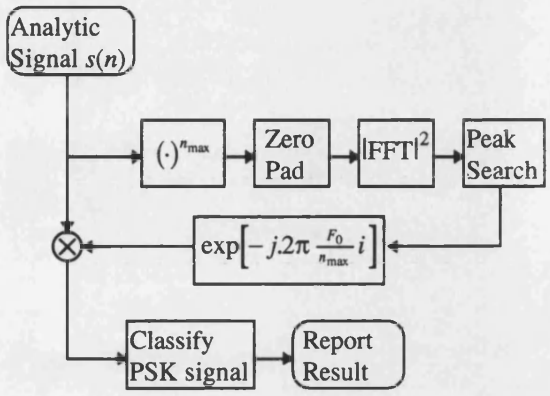


Figure I.1 : Modulation Blind Carrier Recovery Algorithm

This technique for carrier recovery will perform well for the highest PSK type, but will increase the carrier noise level further from the optimum for the other PSK types. The method illustrates a modulation blind PSK carrier synchronisation algorithm, but the performance of the technique is not generally acceptable.

I.2.2 Decision Directed Carrier Recovery

As the detection of the spectral line is worse than the performance of the PSK classifiers, it is quite feasible that misclassification may result from spurious components in the spectrum. The PSK classification methods may be used to confirm the result from a spectral line in the power law.

The technique shown in figure I.2 searches for discrete carrier components at each power of n and attempts classification from these. A level of classification confidence is then attributed to the PSK type corresponding to the power number. This is repeated for each power, and the most probable PSK type is classified.

A method for determining a measure of the classification confidence for maximum likelihood techniques is given in Appendix I.B. Using the maximum likelihood techniques reported in chapter 3, classification was performed on a scaled version of the class probability, as the only information which was sought after was the most likely class. The technique in Appendix I.B describes the estimated probability of a class being correct given the observed data.

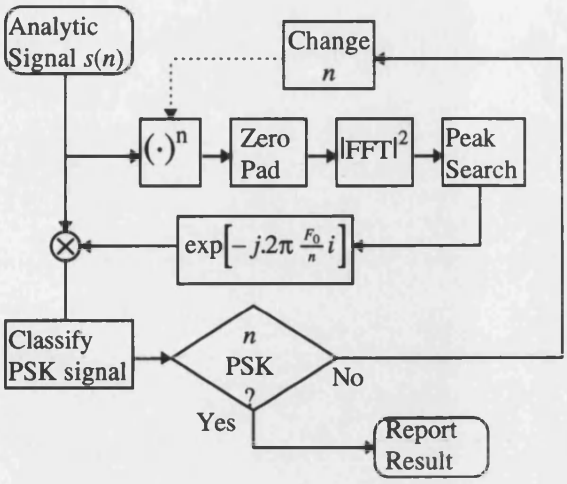


Figure I.2 : Decision Directed Carrier Recovery Algorithm

I.2.3 Other Combinations

Two techniques have been discussed for carrier synchronisation using the power law, and other combinations of processes may be devised for carrier recovery. As an example, the technique in I.2.2 may be base modified to apply the carrier estimate on the three highest peaks.

I.2.4 Iterative Search

An alternative method for overcoming the carrier frequency estimation is to classify the PSK signal whilst sweeping through a range of carrier frequencies around the estimated carrier frequency. At each stage classification should be attempted, and it is assumed that the carrier is correctly placed when a high degree of confidence is obtained in the result.

One such technique is demonstrated in figure I.3. It is assumed that the carrier frequency is approximately known and the signal is mixed down to approximately baseband. Classification is then performed and a decision is made from the classifier as to the significance of the result. If it is deemed that the result does not reflect a high significance, then the signal is mixed to a different frequency and the process is repeated.

The general performance of such an arrangement will be slow due to the number of complex mixing processes which occurs for a range of carrier frequencies.

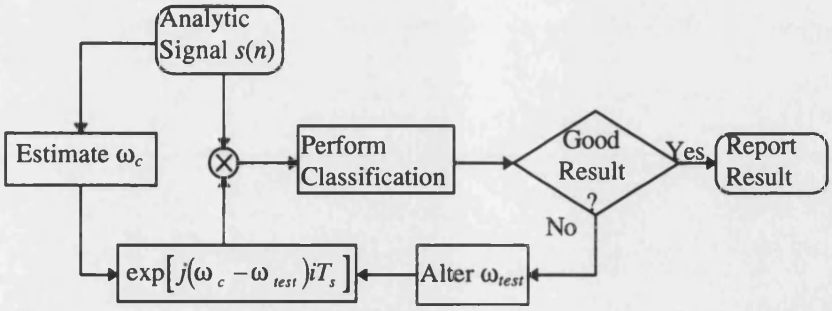


Figure I.3 : Carrier Search Method

The required frequency resolution of such a technique may be calculated from the results in chapter 4 of figures 4.15-4.21. Assuming a transmission of e.g. 64Kb/s and classification based upon 1024 samples, the maximum end of phase offset is 0.25π. This corresponds to a frequency offset of ±8Hz in each step, i.e. the step size may be less than 16Hz.

I.3 Symbol Timing Recovery

The recovery of symbol timing enables the reduction of noise on the detected symbols, and may be performed as a matched filtering process. In the non-coherent techniques for PSK classification, knowledge of the symbol rate and time must be determined so that the phase differencing may be performed on adjacent symbols.

It is not always possible to have knowledge of the symbol characteristics and a general form of filtering must be assumed in such cases. An integrate and dump section is appropriate for a rectangular symbol, and will provide performance improvement for any symbol shape. The general structure is given in figure I.4.

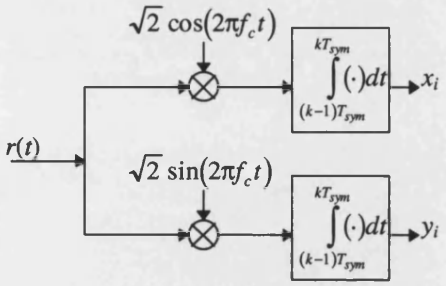


Figure I.4 : Integrate and Dump Filter

This presents an analogue form of matched filter. A digital method would be to take a number of signal samples per symbol, followed by numerical integration on each symbol to approximate the integration process. A simple means for achieving this is to establish the

arithmetic mean of the signal, which approximates to the trapezoidal method for numerical integration.

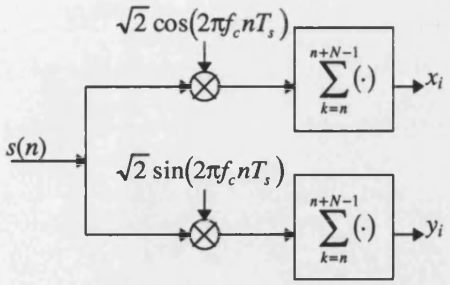


Figure I.5 : Simplified Digital Integrate and Dump Filter

The integration process will average out the effects of noise, and if adjacent samples have a degree of independence, the output of the matched filter will appear more Gaussian in nature. This strengthens the classifier assumption that the noise is attributed to a Gaussian source.

If the symbol pulse shape is known, a similar structure may be applied, where the signal entering the integrate and dump is multiplied by a replica of the pulse shape, or a sampled replica in the digital implementation.

I.3.1 Methods For Timing Recovery

In order to perform the matched filtering of a signal, knowledge is required of the symbol rate and the phase (or timing) of the signal. A number of techniques for achieving this are detailed in [3]. Some useful techniques are detailed in the following discussion, which will produce a reference signal regardless of the number of PSK levels on the signal.

One such technique is to use a delay and multiply type structure [3], where the signal is taken, delayed by half a symbol period and is then multiplied by itself, or a conjugated version in an analytic system. The structure of this is given in figure I.6 where a complex signal is assumed.

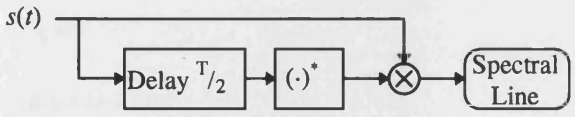


Figure I.6 : Delay and Multiply Detector

Such a technique is useful for the detecting the presence of a digitally modulated signal, as it is valid for different modulation formats, including different QAM, PSK, ASK and FSK modulation formats. An example of this technique is given in figure I.7 where 8PSK is applied at 0dB SNR.

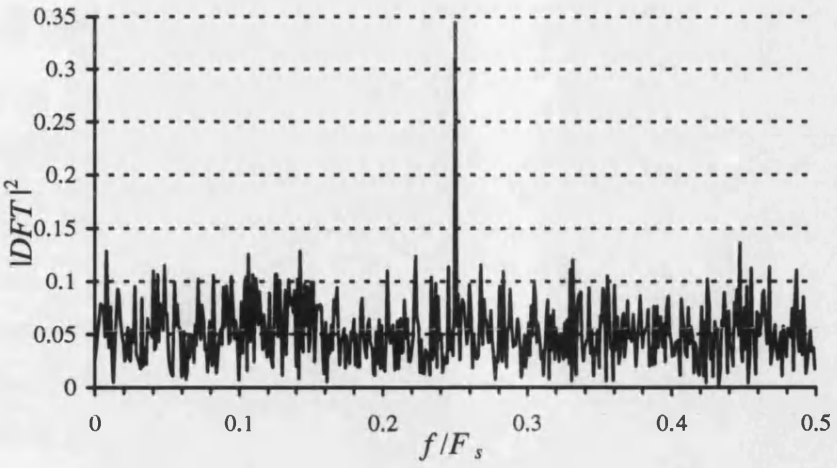


Figure I.7 : Power Spectrum Plot for 8PSK at $F_{sym}=0.25F_s$ at 0dB SNR, 1024 Samples

Another method for generating timing information is to extract a spectral line from the envelope of a band-limited signal. In the case of PSK, the band-limiting of the signal causes the envelope to drop at the symbol transitions, where the instantaneous band-width is large. This then produces a periodic component at the bit rate.

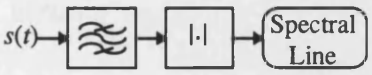


Figure I.8 : Spectrum from the Band-Limited Envelope

A further technique is to extract the symbol information when the signal is close to base band, band-limit the signal and differentiate it. The rectification or other nonlinearities will cause a spectral line to appear at multiples of the bit rate. This method is similar to that of the non-decision directed maximum likelihood timing estimate [3].

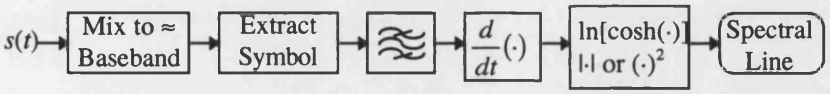


Figure I.9 : Derivative of Symbol with Non-Linearity

This has listed just some of the available methods for symbol synchronisation.

I.4 SNR and Amplitude Estimation

In the parametric classification techniques knowledge of the SNR is assumed. This is a difficult parameter to estimate, especially as the modulation type is unknown.

One method for estimating the noise is to assume that it is all attributed to noise in the receiver, which then enables an estimate to be based on the received power. Many modern receivers enable measurement of the received signal power, which is attributed to the power of the signal plus the noise. The estimate of the signal power and noise power or noise spectral density may then be based upon the noise figure and temperature of the receiver.

Other techniques may assume that the envelope is constant with AGWN present upon it. The SNR is estimated by processing statistics of the observed sampled signal $f(n)$, which is comprised of a wanted signal $s(n)$ and AGWN of standard deviation σ. The following results should first be noted from table 2.1 of chapter 2.

$$\mathrm{E}\left[f(n)f^*(n)\right] = \mathrm{E}\left[s(n)s^*(n)\right] + 2\sigma^2 \tag{I.4}$$

$$\mathrm{E}\left[f(n)^2 f^*(n)^2\right] = \mathrm{E}\left[s(n)^2 s^*(n)^2\right] + 8\sigma^2\,\mathrm{E}\left[s(n)s^*(n)\right] + 8\sigma^2 \tag{I.5}$$

For a constant envelope signal

$$\mathrm{E}\left[s(n)^2 s^*(n)^2\right] = \mathrm{E}\left[s(n)s^*(n)\right]^2 = A^4 \tag{I.6}$$

In order to solve for the SNR ρ, (I.6) is placed in (I.4) and (I.5) to give :

$$\rho = \left[\left(2 - \frac{\mathrm{E}\left[f(n)^2 f^*(n)^2\right]}{\mathrm{E}\left[f(n)f^*(n)\right]^2}\right)^{-\frac{1}{2}} - 1\right]^{-1} \tag{I.7}$$

$$A^2 = \left(\mathrm{E}\left[f(n)^2 f^*(n)^2\right] - 2\,\mathrm{E}\left[f(n)f^*(n)\right]^2\right)^{\frac{1}{2}} \tag{I.8}$$

I.5 Conclusions

This section has looked into techniques for carrier synchronisation, symbol timing and SNR estimation in a modulation recognition environment. Some possible methods have been discussed, but future research is clearly required in these areas.

I.6 References

[1] D.C. Rife, R.R. Boorstyn, "Single-Tone Parameter Estimation from Discrete-Time Observations", IEEE trans. Information Theory, Vol. IT-20, No.5, Sept. 1974, pp. 591-598.

[2] J. Reichert, "Automatic Classification of Communication Signals Using Higher Order Statistics", Proc. ICASSP'92, Vol. 5, pp. V.221-V.224.

[3] J.G. Proakis, "Digital Communications", 3rd Ed., McGraw-Hill, New York, 1995, Ch. 6.

I.7 Appendix I.A

Power Law Classifier

The following development examines the effect of AGWN on a PSK signal when a power law is applied to the analytic signal. The examination is performed on a rectangular symbol in order to establish the greatest performance. When symbol shaping is included, the detectability will become reduced.

Consider an M level PSK signal received in AGWN at a signal to noise ratio ρ. This may be written as :

$$f(i) = A\exp(j\omega_c iT_s)\exp(j\Phi(i)) + n_I(i) + jn_Q(i) \tag{A.1}$$

where $\Phi(i)$ is the PSK phase signal, and includes all the PSK symbol types. n_I and n_Q are Gaussian variables of variance σ^2. When the signal component is taken to an n^{th} power, the following transformation is observed :

$$f_{sig}^n(i) = A^n \exp(jn\omega_c iT_s)\exp(jn\Phi(i)) \tag{A.2}$$

and for M level PSK, where n is an integer multiple of M the phase is represented by :

$$\exp(jn\Phi(i)) = 1 \tag{A.3}$$

which yields a spectral line at n times the carrier frequency. The unwanted signal has a power given by :

$$Var\left[f_{noise}^n\right] = E\left[f^{\frac{n}{2}} f^{*\frac{n}{2}}\right] - A^n \tag{A.4}$$

This resembles the moments of the envelope of a sinusoid in Gaussian noise (Chapter 2, section 2.3.1), and the moments are given in terms of the Confluent Hypergeometric function as :

$$Var\left[f_{noise}^n\right] = (2\sigma)^n \Gamma(n+1)_1 F_1\left(-n;1;-\frac{A^2}{2\sigma^2}\right) \tag{A.5}$$

which may be simplified to

$$Var\left[f_{noise}^n\right] = \left(2\sigma^2\right)^n n!^2 \sum_{i=0}^{n-1} \frac{A^{2i}}{2^i \sigma^{2i} i!^2 (n-i)!} \tag{A.6}$$

Assuming that the spectral line is coincident with a DFT bin (this becomes valid in the general case when zero padding interpolation is used in the DFT), the DFT will consist of a signal bin and a set of noise bins. The magnitude squared DFT is used, where the DFT is scaled by the number of samples. With a large number of bins N, the distribution of the noise bins tends towards a Rayleigh distributed variable, with a mean given by :

$$p_n(y) = \frac{1}{\mu_n}\exp\left(-\frac{y}{\mu_n}\right) \tag{A.7}$$

where :

$$\mu_n = \frac{Var\left[f_{noise}^n\right]}{N} \tag{A.8}$$

The pdf of the signal is given by a non-central Chi squared distributed variable, with two degrees of freedom, given by :

$$p_s(y) = \frac{1}{\mu_n}\exp\left(-\frac{y+A^2}{\mu_n}\right)I_0\left(\frac{2A^n\sqrt{y}}{\mu_n}\right) \tag{A.9}$$

I.7.1.1 Case When n<M

If the tested PSK is not the true PSK, the spectrum of the transformed signal in (A.2) is no longer discrete in nature because the constellation is jumping from point to point. The overall effect is that the spectrum will have a component which is due to the autocorrelation of the digital signal, and this is raised by a noise floor attributed to the interaction of the AGWN.

The DFT spectrum of the noise free signal results from a random process, and the spectrum has a Rayleigh distribution similar to that of noise. The normalised DFT power spectrum of a noise-free base-band PSK signal is given by :

$$S(k) = \frac{A^{2n}T_b}{NT_s}\sin c^2\left(\frac{\left(\frac{2\pi}{N}k - n\omega_c T_s\right)T_b}{2T_s}\right) \tag{A.10}$$

The distribution of the signal and noise on each DFT bin is Rayleigh distributed with mean and standard deviation given by :

$$\mu(k) = \sigma(k) = \frac{A^{2n}}{N}\left[\frac{T_b}{T_s}\mathrm{sinc}^2\left(\frac{\left(\frac{2\pi}{N}k - n\omega_c T_s\right)T_b}{2T_s}\right) + \frac{n!^2}{\rho^n}\sum_{i=0}^{n-1}\frac{\rho^i}{i!^2(n-i)!}\right] \tag{A.11}$$

I.7.1.2 Probability of Misclassifying a Peak

The calculation of the classification performance of the DFT power method may be evaluated for some simple cases. If it is known that a discrete component exists within the spectrum, then the probability of the peak being attributed to the carrier component may be calculated.

It was seen that, for a discrete component the pdf is non-central χ^2 distributed with two degrees of freedom, and the other DFT bins are Rayleigh distributed. The classification performance for a similar problem was evaluated for the DFT of phase histogram (Section 3.4.5). From this, the probability of choosing the wrong peak is given by :

$$P_{err} = \sum_{i=1}^{N-1} \frac{n!(-1)^{i+1}}{(n-i)!(i+1)!} \exp\left(\frac{-i}{i+1}\frac{A^{2n}}{\mu_n}\right) \tag{A.12}$$

From (A.6) it was found that

$$\frac{A^{2n}}{\mu_n} = \frac{N\rho^n}{n!^2 \sum_{i=0}^{n-1} \frac{\rho^i}{i!^2(n-i)!}} \tag{A.13}$$

This is difficult to evaluate for a large number of DFT bins N, as numerical error becomes a problem for N>40. Therefore a simplifying approximation is used, which improves with increasing SNR.

The approximation looks at the probability of N-1 independent events of the signal being in error with one noise bin. The true error considers the signal being in error with N-1 noise bins. This is a subtly different approach which provides poor results at higher error probabilities. However, they are seen to converge for low error probabilities. The probability of error is approximated by :

$$P_{err} = 1 - \left[1 - \frac{1}{2}\exp\left(\frac{-A^{2n}}{2\mu_n}\right)\right]^{N-1} \tag{A.14}$$

The SNR for 1% misclassification probability is described in table. Included in the table I.1 are the 1% misclassification probabilities for the DFT of phase histogram technique. The comparison of these two results must be made with caution.

In practice a peak searching algorithm will not distinguish between the peak from noise and the peak from a signal. Therefore more processing is required to establish the significance of a peak to provide a candidate for a PSK scheme. The performance of a power law classifier will degrade when the peak significance is used.

	CW	BPSK	QPSK	8PSK
DFT power	-16.7dB	-6dB	2.8dB	10.4dB
DFT of Phase Histogram	-19dB	-7dB	1.5dB	8.0dB

Table I.1 : 1% misclassification Probabilities

A further degrading factor occurs when the symbols are not rectangular, as the power in the discrete spectral line becomes reduced, and other lines are also formed at multiples of the bit rate.

It is therefore feasible that decision theoretic techniques may be used to enhance the performance of a power law classifier.

I.8 Appendix I.B

Incorporating a Measure of Confidence in the Maximum Likelihood Classifier

In Appendix 3.B of chapter 3, the maximum likelihood criteria was established using :

$$p(\alpha_m|x_1,x_2\ldots x_N)=p(x_1,x_2\ldots x_N|\alpha_m)\frac{p(\alpha_m)}{p(x_1,x_2\ldots x_N)} \tag{B.1}$$

and it was assumed that each hypothesis is equiprobable and the term $p(x_1,x_2 \ldots x_N)$ is common to all PSK types, and therefore no further use was made of it. However, this term may be of use to determine a meaningful numerical measure for $p(\alpha| x_1,x_2\ldots x_N)$, which can then be used as a metric for classification confidence.

The following may be used to determine the probability of the data :

$$p(x_1,x_2\ldots x_N)=\sum_{m=1}^{C}p(x_1,x_2\ldots x_N|\alpha_m)p(\alpha_m) \tag{B.2}$$

where α_m is the m^{th} hypothesis and C is the number of candidate hypotheses. The maximum likelihood structure may be modified to enable the absolute probability of classification, and the result is shown in figure I.10, where it is assumed that each class is equiprobable.

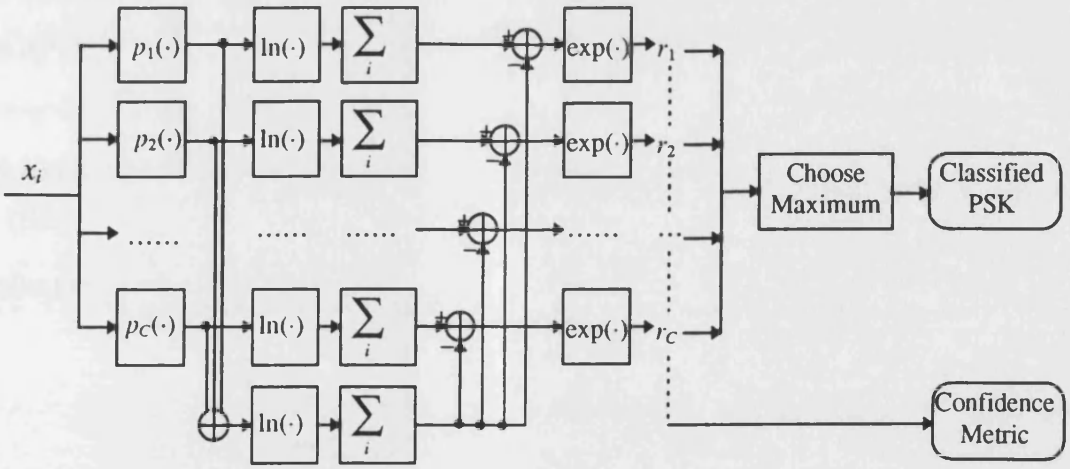


Figure I.10 : Maximum Likelihood Classifier with Confidence Metric

Appendix II : Digital Methods for Analytic Signal Generation

II. Digital Methods for Analytic Signal Generation

II.1 Outline

In order to analyse the characteristics of a signal for modulation recognition, the signal is processed in an analytic form. This enables parameters such as the envelope, instantaneous frequency and phase of the signal to be estimated. This Appendix starts with a definition of the analytic signal and some of the associated properties.

An analytic signal translates a signal into vector form through the Hilbert transform. There are various techniques for achieving the Hilbert transformation and a range of methods have been applied to modulation recognition. Various methods are discussed with respect to implementation and performance properties and this enables a useful comparison between the techniques from a modulation recognition perspective.

In the performance calculations of the Hilbert transform techniques in Appendices A-D it is assumed that the signal is comprises a large number of equal amplitude sinusoids spread across the pass band of the transformer.

II.2 Definition

The definition of an analytic signal is given by [1][2]:

$$z(t) = s(t) + j\,\mathrm{H}[s(t)] \tag{II.1}$$

where H(·) is the Hilbert transform operator. For convenience a Hilbert transformed signal is denoted by :

$$s(t) \xrightarrow{\mathrm{H}(\cdot)} \breve{s}(t) \tag{II.2}$$

The Hilbert transform is essentially a filter of impulse response given by :

$$g(t) = \frac{1}{\pi t} \tag{II.3}$$

and the frequency response of such a filter is given by :

$$G(f)=\begin{cases}-j & f>0\\ 0 & f=0\\ j & f<0\end{cases} \tag{II.4}$$

From (II.4) it can be seen that $|G(f)|=1$ and $\angle G(f)=-\mathrm{sgn}(f)\frac{\pi}{2}$

Some useful identities of the Hilbert transform are related to the transforms of the sinusoidal identities, given by :

$$\mathrm{H}[a(t)\cos(\omega_c t)]=a(t)\sin(\omega_c t) \tag{II.5}$$

$$\mathrm{H}[a(t)\sin(\omega_c t)]=-a(t)\cos(\omega_c t) \tag{II.6}$$

II.3 Analytic Signal Properties

The general form of a modulated signal with amplitude modulation $A(t)$, phase modulation $\Phi(t)$ and angular carrier frequency ω_c is given by :

$$s(t)=A(t)\cos(\omega_c t+\Phi(t)) \tag{II.7}$$

The analytic representation of this signal is denoted by $z(t)$ and is given by (II.1) applied to (II.7) and (II.5) to give :

$$z(t)=A(t)\exp[j(\omega_c t+\Phi(t))] \tag{II.8}$$

In modulation recognition the envelope and phase of the signal are important parameters. The envelope is simply derived using the following property :

$$A(t)=|z(t)|=\sqrt{\mathrm{Re}[z(t)]^2+\mathrm{Im}[z(t)]^2} \tag{II.9}$$

In order to derive the phase $\Phi(t)$ of the signal the carrier must be removed. This may be achieved by multiplying the analytic signal by a complex phasor of opposite frequency :

$$z_0(t)=z(t)e^{-j\omega_c t}=A(t)e^{j\Phi(t)} \tag{II.10}$$

The phase may then be deduced from (II.10) as :

$$\Phi(t)=\angle z_0(t)=\tan^{-1}(\mathrm{Im}[z_0(t)],\mathrm{Re}[z_0(t)]) \tag{II.11}$$

where the arctangent function includes all four quadrants of phase.

The power spectral density of a Hilbert transformed signal is the same as that of the original signal [2], but the spectrum of an analytic signal is given by :

$$\Im(z(t)) = \begin{cases} 2S(\omega) & \omega > 0 \\ S(\omega) & \omega = 0 \\ 0 & \omega < 0 \end{cases} \qquad \text{(II.12)}$$

II.4 Implementation

There are a number of methods for achieving the Hilbert transformation, and four of the more practical methods are discussed in the following sections.

II.4.1 Finite Impulse Response Filter Implementation

This method uses an FIR filter to perform the Hilbert transformation. The general structure of the FIR filter uses an odd number of coefficients and odd symmetry on the filter coefficients. This odd symmetry provides an exact 90° phase shift, but does not provide a perfect magnitude response.

The analytic structure of the FIR method is provided in figure II.1. The delay is included to match the signal to the group delay of the filter and thus avoid a phase imbalance between the two channels. For this reason the FIR filter must contain an odd number of taps in order to avoid a half sample delay.

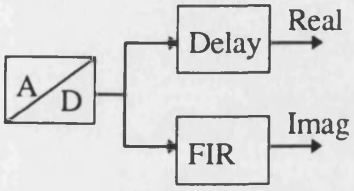


Figure II.1 : FIR Analytic Signal Generation

The equation for the generation of the Hilbert transformed components is given by :

$$y(n) = \sum_{r=1}^{N} h(r)[x(n+r) - x(n-r)] \qquad \text{(II.13)}$$

where $h(n)$ is the impulse response of the filter, $n \in \{-N,-N+1,\ldots N\}$, and the filter therefore contains $2N+1$ coefficients.

The FIR filter may be designed using a variety of methods, but the Parks-McClellan approach [4] provides an iterative method to create a magnitude response which has equal ripple in the

pass-band. This is a desirable response in many applications as the error is uniformly distributed across the pass-band.

The response of a practical FIR Hilbert transformer is described solely in terms of the amplitude response, as the phase response is always perfect. Using a Parks McClellan design the design parameters will specify the pass band ripple and the bandwidth of the pass band. The amplitude response drops to zero at D.C. and the Nyquist frequency, to give a set of design parameters characterised in figure II.2.

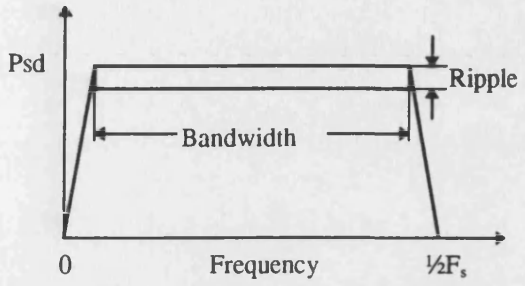


Figure II.2 : FIR Hilbert Transform Design

A desirable filter design will have maximum bandwidth and minimum pass-band ripple. The bandwidth parameter determines the available frequency band that a signal may occupy. The deviation due to the pass-band ripple causes unwanted signal components to be produced in the analytic signal.

The number of filter taps is traded off with the ripple and bandwidth of the filter. In order to illustrate this trade-off, two designs are illustrated in figure II.3. The first has 23 taps, ±0.1dB ripple and a pass band of $0.4F_s$. The second has 71 taps, ±0.01dB ripple and a pass band of $0.45F_s$.

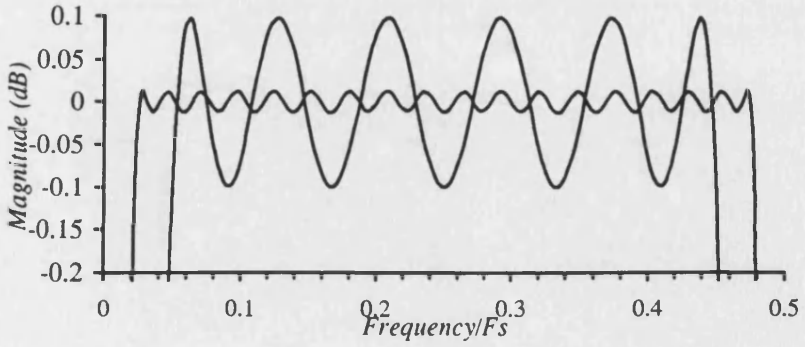


Figure II.3 : Amplitude Response of 23 tap and 71 tap Hilbert transformers.

The above responses are zero at the odd points of the impulse response. This is achieved when $\frac{N-1}{2}$ is an odd integer and enables a more computationally efficient solution, requiring half the number of multiplications and additions that would otherwise be required.

$$y(n) = \sum_{r=0}^{\frac{N-1}{2}} h(2r+1)\left[x(n+2r+1) - x(n-2r-1)\right] \tag{II.14}$$

The rejection ratio for the wanted to unwanted power has been calculated in Appendix II.A and is given by :

$$\gamma_{db} \approx 25 - 20\log_{10}(\alpha) \tag{II.15}$$

where $\pm\alpha$ is the pass-band ripple. In the 23 tap transformer, the rejection ratio is approximately 45dB, and the 71 tap transformer, the rejection ratio is approximately 65dB.

II.4.2 Discrete Fourier Transform method

This technique uses frequency domain methods to achieve the Hilbert transform result in (II.4), and the process is described in figure II.4.

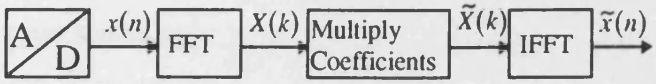


Figure II.4 Generation of the Hilbert Transform Using the FFT

$$\tilde{X}(k) = \begin{cases} 0 & k = 0, \frac{N}{2} \\ -jX(k) & k = 1 \ldots \frac{N}{2} - 1 \\ jX(k) & k = \frac{N}{2} + 1 \ldots N - 1 \end{cases} \tag{II.16}$$

The DFT coefficient multiplication is described in (II.16), where it can be seen that the appropriate $\pi/2$ phase shifts are introduced through the complex operator j.

The deviation of the DFT method from the ideal Hilbert transformer is different to that of the FIR Hilbert transformer. Using the FIR method the error was due to a linear amplitude imbalance at a frequency. Using the FFT method the signal becomes distorted in the time domain and the amount of distortion is related to the applied frequency.

This distortion occurs at frequencies which are not coincidental with the FFT bins, and the distortion is time varying, with maximum error at the ends of the transformation and minimum error at the centre.

An example of the distortion effect is provided in figure II.5 where a plot of error modulus against sample number is provided for a 64 point transform of a 64 point unity sinusoid at a frequency of $\frac{5.5}{64}F_s$. The modulus is used for clarity, but in practice the error changes sign every sample.

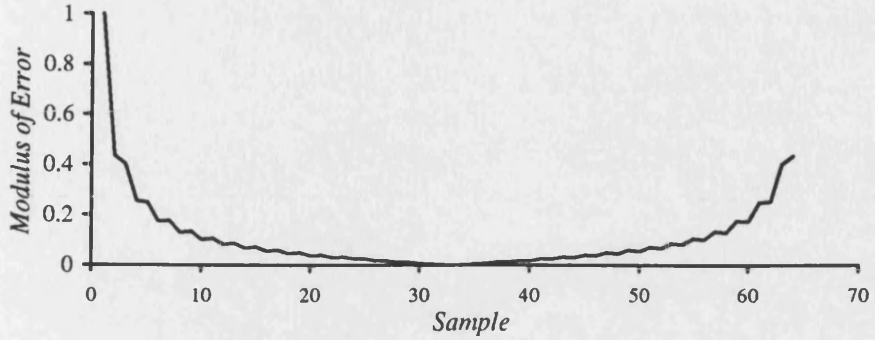


Figure II.5 : A Plot of Error Modulus against Sample Number for a 64 point DFT Hilbert Transform

The frequency domain effects are observed in figure II.6, where the distortion error generally increases as the frequency moves either side of $\frac{1}{4}F_s$. Figure II.6 illustrates the normalised mean square error for a DFT Hilbert transformed sine wave against frequency for a 32 and a 256 point process. It is seen from the plots that local maximum errors occur at frequencies mid-way between frequency bins.

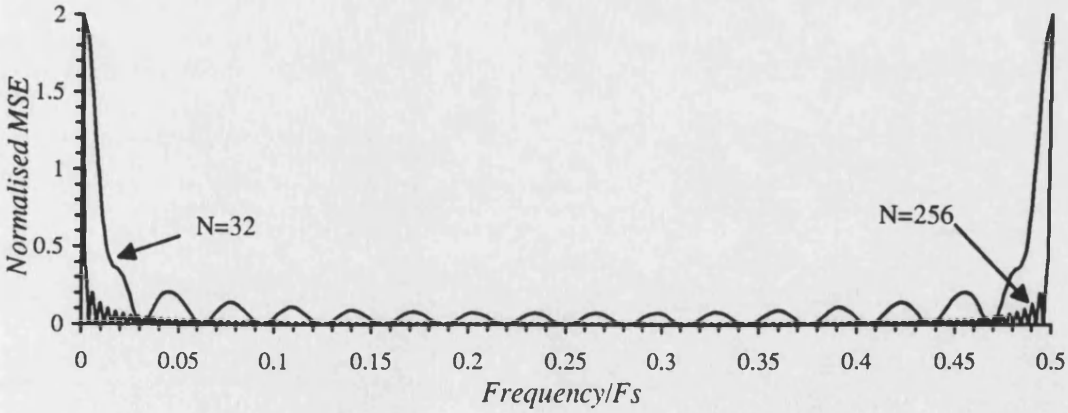


Figure II.6 : Plots of Hilbert transform Mean Square Error against Frequency for the DFT Method

Figure II.6 also shows that as the number of bins are increased, the number of maxima increases as would be expected, but the peaks are similar in magnitude when the frequency scales are normalised. Figure II.7 shows this more clearly where the two examples have the frequency scales warped to place the nulls in identical places.

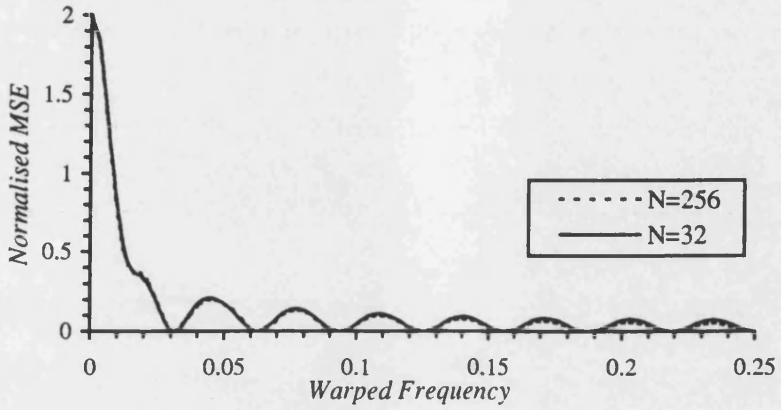


Figure II.7 : Comparison of Mean Square Error of Localised Peaks for N=32 and N=256

As the number of samples entered into the Hilbert transform is increased, the distorted components tend to become confined towards the zero frequency and Nyquist bands, thus providing a more desirable characteristic. This trend is similar to an increase in the number of taps used in the FIR technique, except that the FIR technique is not subject to time domain distortion.

As the number of samples is increased the number of floating point operations increases. The execution time of the FFT T_{exec} for N points is roughly given by the following relationship [5][6]:

$$T_{exec} = \mu N \log_2(N) \tag{II.17}$$

where μ is a processor related constant. When the number of samples is increased by a factor k, the execution time T'_{exec} increases to :

$$T'_{exec} = kT_{exec} + \mu Nk \log_2(k) \tag{II.18}$$

i.e. the processing time is larger than k individual blocks. The number of samples which are processed in each block therefore has a direct impact on the overall system computational complexity.

The wanted to unwanted signal power ratio has been derived in Appendix II.D, and a graph is provided in figure II.14 of the appendix. For a 1024 sample transformation the rejection ratio was found to be approximately 30dB.

The DFT method is a simple technique for obtaining the Hilbert transform of a signal and the FFT of the signal is calculated as an intermediate process in many modulation recognition applications. The main drawback of the technique is the distortion involved in the process, producing unwanted signal components in the analytic signal. Another drawback is the delay involved when a large number of samples are processed. This is however not a major drawback in modulation recognition applications as the samples are generally post processed.

II.4.3 Quadrature Mixing method (DSP)

The sampled signal may be converted into analytic form by sampling the signal at an IF and then by mixing the signal by two quadrature components followed by low pass filtering. A block diagram of this is provided in figure II.8.

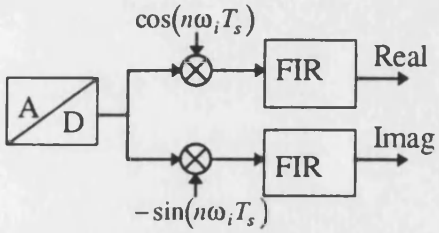


Figure II.8 : Quadrature Method for Generating an Analytic Signal

The digitised signal may be represented by :

$$s(n) = A(t)\cos(n\omega_i T_s + \Phi(n)) \tag{II.19}$$

and the two parallel processes in figure II.8 are as follows :

$$s(n)\cos(n\omega_i T_s) = \tfrac{1}{2} A(t)\cos(\Phi(n)) + \tfrac{1}{2} A(t)\cos(2n\omega_i T_s + \Phi(n)) \tag{II.20}$$

$$-s(n)\sin(n\omega_i T_s) = \tfrac{1}{2} A(t)\sin(\Phi(n)) - \tfrac{1}{2} A(t)\sin(2n\omega_i T_s + \Phi(n)) \tag{II.21}$$

After low pass filtering, the two signals have been translated down in frequency by ω_i and satisfy the conditions of an analytic signal. In order to maximise the bandwidth of the signal, ω_i is set to approximately a quarter of the sampling rate. If it is set exactly at a quarter of the

sampling rate, the generation of the sine and cosine coefficients becomes a set of repeated constants where one period is described below :

$$\cos\left(2\pi n \frac{F_s}{4} T_s\right) = \left\{1, \frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}, -1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, \ldots\ldots\right\} \quad \text{(II.22)}$$

$$-\sin\left(2\pi n \frac{F_s}{4} T_s\right) = \left\{0, -\frac{1}{\sqrt{2}}, -1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 1, \frac{1}{\sqrt{2}}, \ldots\ldots\right\} \quad \text{(II.23)}$$

Figure II.9 illustrates the transformation in signal spectrum in the frequency band ±½*Fs*. In (a) the spectrum of the incoming signal is displayed and (b) shows the spectrum after the multiplication by the quadrature sinusoids when the two components are considered in analytic form. The dashed line represents the desired filter response.

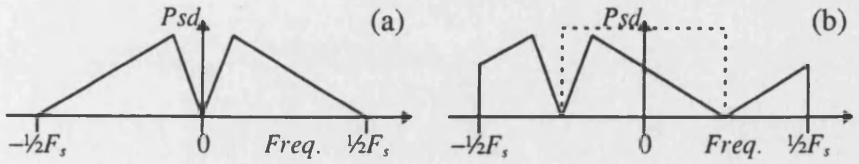


Figure II.9 : (a) Spectrum of the real input signal (b) Spectrum of the Analytic Signal Prior to Filtering

The filter may be implemented using an FIR filter which preserves linear phase, and again the Parks-McClellan equiripple filter design [4] may be applied. The number of filter taps increases when the pass-band ripple is decreased, stop-band attenuation is increased and the transition between pass and stop band is decreased (Ref. Figure II.10). The design is therefore a compromise between computational complexity and performance.

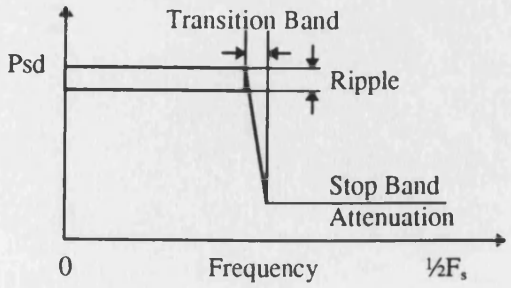


Figure II.10 : FIR Filter Parameters

Two examples of filter designs for such a task are shown in figure II.11. The 71 tap filter has ±0.1dB ripple and 70dB stop-band attenuation. The 23 tap filter has ±0.3dB ripple and 60dB stop-band attenuation. This provides an indication of trade between the number of filter taps and the transition bandwidth which may be attained.

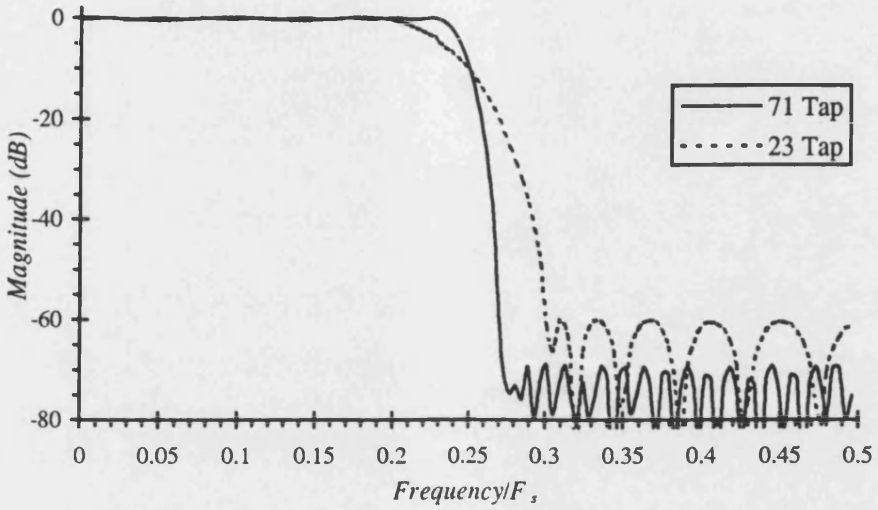


Figure II.11 : FIR Filter Designs for Image Frequency Rejection for 23 and 71 Tap Filters

This method for Hilbert transforming the signal is a useful one, which balances out the response in the in-phase and quadrature channels due to two identical FIR filters, whereas the FFT method and FIR Hilbert transform method will experience an imbalance due to the distortion and non-ideal frequency responses. However, the method is restricted by the stop-band and pass-band characteristics of the filter which are a function of computational complexity.

Following the frequency translation the signal may be decimated by a factor of two to reduce the sampling frequency of the signal. This is illustrated in figure II.12 which follows from figure II.9 (b), where the filtered signal is decimated.

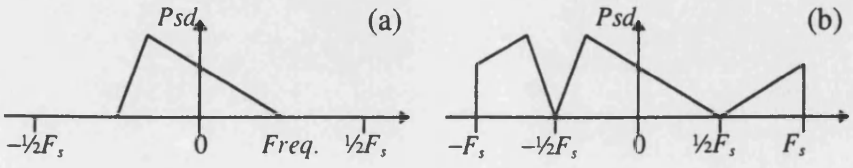


Figure II.12 : (a) Signal after translation (b) The Decimation Process

The method is computationally intensive when compared to the Hilbert transform FIR method, as the technique requires two filtering operations along with a frequency translation. Also the Parks-McClellan type of filter design algorithm for the low-pass filter case does not lend itself to the possibility of creating every odd sample with a zero like the Hilbert transformer design.

Appendix II.C derives the ratio of wanted to unwanted signal power, and it is approximated by :

$$\gamma_{dB} \approx 3 - \beta \qquad \text{(II.24)}$$

Using the above filter designs as example parameters, the 23 tap filter yields a 63dB rejection ratio, and the 71 tap filter has a 73dB rejection ratio.

There are commercially available solutions which tackle the problem in this manner, where the signal is digitally sampled at a high I.F., filtered, translated in frequency, decimation filtered and is then decimated. Such a solution can be realised with the hardware in [7] which acts as a parallel DSP, and provides a means for using the full bandwidth range of the signal as well as enabling an adaptive set of sampling frequencies.

II.4.4 Quadrature receiver

The method in II.4.3 performs the Hilbert transform after the sampling process. However, it is possible to perform this prior to the sampling, as shown in figure II.13.

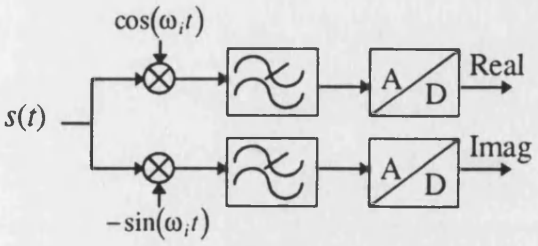


Figure II.13 : Quadrature Receiver Implementation of an Analytic Signal

The theory behind this technique is identical to that of the DSP quadrature receiver in II.4.3, but there is a substantial difference in the implementation. The receiver takes the signal from the radio frequency or an I.F. and mixes it separately with two orthogonal carriers. One component will be at twice the carrier frequency and the other will be at an I.F. of approximately zero. The twice frequency component is then filtered out by the anti-alias filter and the signal at the wanted frequency is digitally sampled.

Two A/D converters are required for this method, whereas the other techniques require only one. The main advantages of this technique is that the signal may be taken close to zero I.F. in one operation, and the bandwidth of the signal may span the Nyquist range.

A final benefit of this technique is that it requires little to no computational overhead in order to evaluate the analytic signal, as all the processing is performed in the analogue sections.

The filters which are used require similar characteristics, otherwise there will not be a cancellation of the negative frequency terms associated with real signals. The phase between the two oscillators must be accurately set to provide a $\pi/2$ phase shift and the amplitude of this source must be balanced.

Some typical amplitude and phase imbalance figures are given in [8] [9] [10], and some typical values of amplitude imbalance are 0.2dB and phase mismatch of 1°. The rejection ratio is derived in Appendix II.B as :

$$\gamma_{dB} \approx 3 - 10\log_{10}\left(\varepsilon^2 + \theta^2\right) \qquad \text{(II.25)}$$

Using these figures, a rejection ratio of 37dB is obtained.

II.5 Conclusions

Four solutions have been analysed for the conversion of a signal into an analytic form, each with relative merits and disadvantages which have been discussed in the preceding sections.

It has been found that the Quadrature DSP and the Hilbert transform methods have a good rejection ratio performance. The FFT method and Quadrature Receiver techniques have both been found to have relatively poor rejection ratios.

A choice of solution for the modulation recognition process may be efficiently realised by a chip set solution similar to [7], which uses the method discussed in II.4.3 for analytic signal generation using DSP quadrature mixing. The sampling frequency of such a chip set can be high and decimation orders higher than 2:1 may be used to select a signal of interest within a band of signals.

The advantages of this solution are that the sampling frequency may be adaptively set, with appropriate anti-aliasing filtering. Signals within a wide bandwidth may then be analysed prior to the selection of an appropriate frequency band, which is useful in the surveillance aspect of modulation recognition. Finally, the chip sets represent a parallel processing technique, which enables additional processing performance within the modulation recognition DSP.

II.6 References

[1] J.G. Proakis, "Digital Communications", 3rd Ed., McGraw-Hill, New York, 1995, Ch. 4.

[2] A. D. Whalen, "Detection of Signals in Noise", Academic Press, New York, 1971, Ch. 3.

[3] A.V. Oppenheim, R.W. Schafer, "Digital Signal Processing", Prentice Hall, London, 1975, Ch. 7.

[4] L. R. Rabiner, J.H. McClellan, T.W. Parks, "FIR Digital Filter Design Using Weighted Chebychev Approximation", Proc. IEEE, Vol. 63, pp. 595-610, April 1975.

[5] "WE DSP32 and DSP32C Application Software Library", AT&T, 1991, Appendix E.

[6] J.D. Martin, "Signals & Processes, a Foundation Course", Pitman, London, 1991, p. 268.

[7] Harris Semiconductor Data Sheets for "HSP45116" and "HSP43220", 1994.

[8] Maxim Integrated Products Data Sheet for "MAX2451".

[9] Harris Semiconductor Data Sheet for "HFA3724".

[10] Mini Circuits Data Book, "MIQ-D" range.

II.7 Appendix II.A

Calculation of the Rejection Ratio for the FIR Hilbert Transform Method

The error in the Hilbert transformation causes unwanted signal components to be placed into the pass-band of the analytic signal. The following derivation will obtain an approximate expression for the ratio of wanted signal to unwanted signal power based on a Parks-McClellan design of Hilbert transformer.

The component of signal at frequency f, $z_f(n)$, is given by :

$$z_f(n) = A(n)\exp\left[j2\pi fnT_s\right] + j\delta(f)A(n)\sin\left(2\pi fnT_s\right) \tag{A.1}$$

which contains a signal and an unwanted component. The unwanted power is due to the frequency dependent pass-band ripple given by $\delta(f)$. The FIR filter is kept to an equal ripple in the pass-band with the Parks-McClellan algorithm, and the ripple approximates a sinusoid when expressed in a logarithm (dB) scale. This may be modelled as :

$$p_{db}(f) = -\alpha\cos\left([N+1]\pi f\right) \tag{A.2}$$

where N is the number of filter taps, and the passband ripple is expressed as $\pm\alpha$ dB. The actual gain of this signal is given by :

$$p_{lin}(f) = \exp\left[-\tfrac{\ln(10)}{20}\alpha\cos\left([N+1]\pi f\right)\right] \tag{A.3}$$

In general the ripple is small, and the expression in (A.3) may be approximated by :

$$p_{lin}(f) \approx 1 - \tfrac{\ln(10)}{20}\alpha\cos\left([N+1]\pi f\right) \tag{A.4}$$

$\delta(f)$ may be deduced from (A.4) and (A.1) by :

$$\delta(f) = -\tfrac{\ln(10)}{20}\alpha\cos\left([N+1]\pi f\right) \tag{A.5}$$

Assuming that the signal is comprised of a large number of sinusoids across the filter pass-band, the average ratio of the mean power of the unwanted signal to the power of the wanted signal is deduced from (A.1) and (A.5) as :

$$\gamma = \tfrac{1}{2}E\left[\delta^2(f), f\right] \approx \tfrac{1}{4}\left[\tfrac{\ln(10)}{20}\alpha\right]^2 \tag{A.6}$$

The ratio of signal power to unwanted signal power ratio expressed in a dB scale is given by :

$$\gamma_{db} \approx 25 - 20\log_{10}(\alpha) \tag{A.7}$$

II.8 Appendix II.B

Calculation of Rejection Ratio for the Quadrature Receiver Method

Assuming that the phase offset between channels is θ radians and the fractional amplitude error is given by ε, the signal $s(t)$ at the output of the mixer is given by :

$$z(t) = s(t)[1-\varepsilon]\cos(\omega_i t + \theta) - js(t)\sin(\omega_i t) \tag{B.1}$$

This may be expanded to :

$$z(t) = s(t)[1-\varepsilon]\left[\cos(\omega_i t)\cos(\theta) - \sin(\omega_i t)\sin(\theta)\right] - js(t)\sin(\omega_i t) \tag{B.2}$$

The signal may be decomposed into a baseband form and a twice frequency component as :

$$s(t)\cos(\omega_i t) = \tfrac{1}{2}s_b(t) + \tfrac{1}{2}s_{2\omega}(t) \tag{B.3}$$

$$-s(t)\sin(\omega_i t) = \tfrac{1}{2}\tilde{s}_b(t) - \tfrac{1}{2}\tilde{s}_{2\omega}(t) \tag{B.4}$$

Therefore at the output of the low pass filter, the signal is given by :

$$z(t) = \tfrac{1}{2}[1-\varepsilon]\left[s(t)\cos(\theta) + \tilde{s}(t)\sin(\theta)\right] + j\tfrac{1}{2}\tilde{s}(t) \tag{B.5}$$

Using approximations for small θ, the expression simplifies to :

$$z(t) \approx \underbrace{\tfrac{1}{2}\left[s(t) + j\tilde{s}(t)\right]}_{\text{Wanted}} - \underbrace{\tfrac{1}{2}\varepsilon\left[s(t) + \tilde{s}(t)\theta\right] + \tfrac{1}{2}\tilde{s}(t)\theta}_{\text{Unwanted}} \tag{B.6}$$

Using the approximation that ε is small, the ratio of wanted to unwanted power is given by :

$$\gamma \approx \frac{2}{\varepsilon^2 + \theta^2} \tag{B.7}$$

Expressed in dB form, this yields :

$$\gamma_{dB} \approx 3 - 10\log_{10}\left(\varepsilon^2 + \theta^2\right) \tag{B.8}$$

II.9 Appendix II.C

Calculation of Rejection Ratio for the Quadrature Mixing Method (DSP)

In this case the I and Q channels are subject to the same pass band ripple, and there is no aliasing in the wanted band. However, there are components which are present at frequencies outside the wanted band, and when decimation is performed this unwanted signal will be in the wanted band. The following work describes the ratio of wanted to unwanted power assuming the signal consists of a large number of sinusoids uniformly distributed in frequency about the pass-band.

The magnitude response of a Parks-McClellan filter in the stop band may be approximated by:

$$|H(\omega)| \approx 10^{-\frac{\beta}{20}} \left|\cos(\pi f(N+1)+\theta)\right| \qquad \text{(C.1)}$$

where β is the stop-band attenuation (dB), N is the number of filter taps and θ is a constant phase offset. When a sinusoidal signal of amplitude A is present in the pass-band, it will become aliased in both the I and Q channels during the decimation process, and the unwanted signal is given by :

$$s(t,f) = 10^{-\frac{\beta}{20}} \left|\cos(\pi f(N+1)+\theta)\right| \frac{A}{2}\cos(2\pi ft) \qquad \text{(C.2)}$$

The average unwanted signal power in a channel is given by :

$$p = \mathrm{E}\left[s^2(t,f); t, f\right] = \frac{A^2}{8} 10^{-\frac{\beta}{10}} \qquad \text{(C.3)}$$

The wanted power to unwanted power ratio γ is given by :

$$\gamma = 2\left(10^{-\frac{\beta}{10}}\right) \qquad \text{(C.4)}$$

and may be expressed in dB form as :

$$\gamma_{dB} \approx 3 - \beta \qquad \text{(C.5)}$$

II.10 Appendix II.D

Calculation of Rejection Ratio for the FFT Method

The normalised mean square error has been calculated against frequency for two cases in figure II.6, which is the ratio of the error signal to the signal power. It will be assumed that the frequency band of interest spans from the first null to the last null in the mean square error, i.e.

$$f \in \left\{ \tfrac{1}{N} F_s \ldots \tfrac{N-2}{2N} F_s \right\} \tag{D.1}$$

The mean square error in figure II.6 is the power of the unwanted signal, and the overall power of the unwanted signal when the input signal comprised of a large number of sinusoids. This is given by :

$$p_{un} = \int_{\frac{F_s}{N}}^{F_s \frac{N-2}{2N}} \varepsilon(f) df \tag{D.2}$$

Note that the normalisation is with respect to the power of a single sinusoid, i.e. divided by ½. The overall wanted signal power to unwanted signal power ratio is given by :

$$\gamma_{db} = -10 \log_{10} \left(\frac{1}{2} \int_{\frac{F_s}{N}}^{F_s \frac{N-2}{2N}} g(f) df \right) \tag{D.3}$$

where $g(f)$ is the response given in the graph in figure II.6. This integral has been evaluated numerically, and results are shown in the graph of figure II.14.

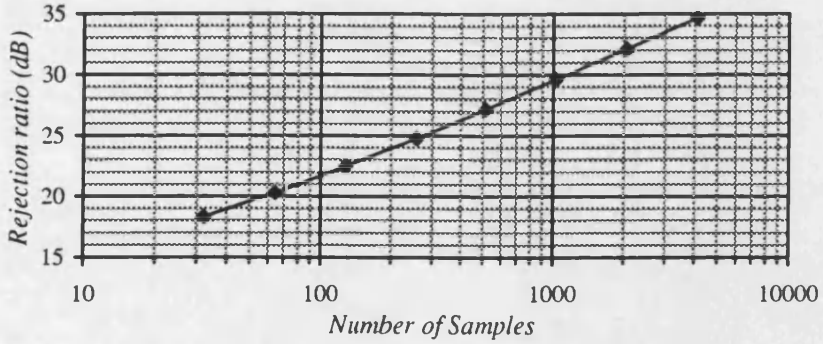


Figure II.14 : A Plot of Rejection Ratio Against The Number of Samples for the FFT Method

From this graph an empirical model may be deduced for determining the rejection ratio :

$$\gamma_{db} \approx 7.85 \log_{10}(N) + 5.93 \tag{D.4}$$

where N is the number of FFT points.

Appendix III

Publications Which Have Arisen From This Research

1. P.C. Sapiano, R.J. Holbeche, J.D. Martin, "Low SNR Approximation To Phase PDF For PSK Signals", IEE Electronics Letters, Vol. 30, No. 16, pp. 1279-1280, 1994. **III.1-2**

2. P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Further Results In The Classification Of PSK Signals Using The Optimum Method", IEE Electronics Letters, Vol. 31, No. 1, pp. 19-20, 1995. **III.3-4**

3. P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Classification of PSK Signals using the DFT of Phase Histogram", Proc. IEEE ICASSP-95, Vol. 3, 1995, pp. 1868-1871. **III.5-8**

4. P.C. Sapiano, J.D. Martin, "Identification of PSK Signals", Proc. IEE Radio Receivers and Associated Systems Conference, No. 415, pp. 95-99, September 1995 **III.9-13**

5. P.C. Sapiano, J.D. Martin, , "Maximum Likelihood PSK Classification using the DFT of Phase Histogram", Proc. IEEE GLOBECOM '95, vol. 2, pp 1029-1033 **III.14-18**

6. P.C. Sapiano, J.D. Martin, "Statistical Performance Of The First Order Phase Difference Digital Instantaneous Frequency Estimator", IEE Electronics Letters, Vol. 32, No. 18, pp. 1657-1658. **III.19-20**

7. P.C. Sapiano, J.D. Martin, "Maximum Likelihood PSK Classifier", Proc. IEEE MILCOM '96, vol. 3, pp. 1010-1014. **III.21-25**

Low SNR approximation to phase PDF for PSK signals

P.C. Sapiano, R.J. Holbeche and J.D. Martin

Indexing terms: Phase shift keying, Statistics for communications, Probability

The authors show that for a low SNR the probability density function of phase for multilevel PSK signals may be approximated as a mean term plus a number of sinusoid terms. A relationship between the SNR and the accuracy of the approximation is developed.

Introduction: The analysis of PSK signals for signal detection often relies on the analysis of the probability density function of the signal under additive Gaussian white noise (AWGN). This analysis is often cumbersome, as the PDF is expressed using an extension of the classic phase distribution of a sinusoidal signal under noisy conditions [1].

$$f(\phi) = \frac{1}{2\pi}e^{-\rho} + \frac{e^{-\rho}}{2}\sqrt{\frac{\rho}{\pi}}\cos\phi e^{\rho\cos^2\phi}\left(1 + \mathrm{erf}\left[\sqrt{\rho}\cos\phi\right]\right) \tag{1}$$

Where ϕ is the phase perturbation about zero and ρ is the SNR. This is extended to L level PSK signals by

$$f_L(\phi) = \frac{1}{L}\sum_{k=0}^{L-1} f\left(\phi + \frac{2\pi(k+0.5)}{L} - \pi\right) \tag{2}$$

Approximations of the phase PDF have been made for a high SNR range using the Tikhonov PDF [2], which is a good approximation for $\rho > 7$dB, while for a larger SNR this may be taken to be a Gaussian PDF. The following reasoning will develop a model for the PDF at low SNR.

Analysis of a sinusoid: As the phase PDF is defined between $[-\pi, \pi]$ the function may be expressed over the full range as a Fourier series with period 2π.

$$f(\phi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{m=1}^{\infty} b_m\cos(m\phi) \tag{3}$$

where

$$b_m = \int_{-\pi}^{\pi} f(\phi)\cos(m\phi)d\phi \tag{4}$$

The integrand is split up into the two parts shown in eqn. 1, and it can be shown that

$$\begin{aligned}&\cos(\phi)\cos(m\phi)\exp[\rho\cos^2\phi]\mathrm{erf}[\sqrt{\rho}\cos\phi]\\ &= \frac{\cos(m\phi)}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{2^{2(n+1)}\rho^{n+\frac{1}{2}}(n+1)!}{(2n+2)!}[\cos\phi]^{2n+2}\end{aligned} \tag{5}$$

The $(\cos\phi)^{2n+2}$ term may be decomposed into a Fourier series using [4], and the product of this and $\cos(m\phi)$ form a set of orthogonal functions which integrate to zero except for one term when m is even. This yields

$$\begin{aligned}b_m &= \rho e^{-\rho}\sum_{n=\frac{m}{2}-1}^{\infty}\frac{(n+1)!}{(n+1-\frac{m}{2})!(n+1+\frac{m}{2})!}\rho^n\\ &= e^{-\rho}\rho^{\frac{m}{2}}\sum_{k=0}^{\infty}\frac{\Gamma(k+\frac{m}{2}+1)}{k!\Gamma(k+m+1)}\rho^k\end{aligned} \tag{6}$$

The other part of the integral is determined using the following series representation:

$$\cos(\phi)\cos(m\phi)e^{\rho\cos^2\phi} = \sum_{n=0}^{\infty}\frac{\rho^n\cos m\phi(\cos\phi)^{2n+1}}{n!} \tag{7}$$

In a similar way to the above it can be shown that this integral is identical to eqn. 6 for odd m.

Therefore it is deduced that this expression is valid for all integer m. The integrand may be expanded as

$$\begin{aligned}&e^{\rho\cos^2\phi}\cos(m\phi)\cos(\phi)\\ &= \frac{1}{2}e^{\frac{\rho}{2}}\exp\left[\frac{\rho}{2}\cos(2\phi)\right]\left(\cos(m+1)\phi + \cos(m-1)o\right)\end{aligned} \tag{8}$$

which may be integrated in terms of modified Bessel functions to yield the result

$$b_m = \frac{\sqrt{\rho\pi}e^{-\frac{\rho}{2}}}{2}\left[I_{\frac{m+1}{2}}\left(\frac{\rho}{2}\right) + I_{\frac{m-1}{2}}\left(\frac{\rho}{2}\right)\right] \tag{9}$$

From the previous argument this may be extended to include odd and even m.

Multilevel PSK: As multilevel PSK is the sum of a number of phase shifted versions of the sinusoidal PDF as in eqn. 2, all the sinusoids will be cancelled except when $m = nL$. This is described as:

$$f_L(\phi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^n b_{(nL)}\cos(nL\phi) \quad (10)$$

As L is even, only even values of m are used in the expression for b_m. Therefore the expression is expressed in terms of modified Bessel functions order $n + 1/2$. This form of function is related to the spherical Bessel function and has recurrence properties [4] that enable the function to be calculated quickly and efficiently.

$$I_{(n+\frac{1}{2}-1)}(x) = I_{(n+\frac{1}{2}+1)}(x) + \frac{2n+1}{x} I_{(n+\frac{1}{2})}(x) \quad (11)$$

For this type of function the recurrence is upwardly unstable, therefore reverse recurrence using the Miller method may be used. This result may then be normalised about $I_{1/2}$ to give the correct result, where

$$I_{\frac{1}{2}}\left(\frac{\rho}{2}\right) = \frac{2}{\sqrt{\pi\rho}}\sinh\left(\frac{\rho}{2}\right) \quad (12)$$

However a further simplification may be introduced. By normalising about $(1-e^{-\rho})/2$ a scaled set of Bessel functions $\ell_v(x)$ are formed where

$$\ell_v(x) = \frac{\sqrt{\pi\rho}}{2} e^{-\frac{\rho}{2}} I_v(x) \quad (13)$$

and

$$b_m = \ell_{\frac{m+1}{2}}\left(\frac{\rho}{2}\right) + \ell_{\frac{m-1}{2}}\left(\frac{\rho}{2}\right) \quad (14)$$

From eqn. 9 it can be shown that

$$b_{m+2} = b_m - \sqrt{\frac{\pi}{\rho}}(m+1)e^{-\frac{\rho}{2}} I_{\frac{m+2}{2}}\left(\frac{\rho}{2}\right) \quad (15)$$

It is evident from eqn. 15 that b_m decreases as m increases. For a PSK signal with large E_s there are fewer harmonics required as the amplitude ratio between the harmonics is larger. It is also found that the spacing between b_m terms increases as the SNR is decreased, therefore the accuracy of the approximation improves with decreasing SNR. The error incurred in a T term approximation is given by

$$\varepsilon(\phi) = f_L(\phi) - \frac{1}{2\pi} - \frac{1}{\pi}\sum_{n=1}^{T}(-1)^n b_{(nL)}\cos(nL\phi) \quad (16)$$

This may be expressed in mean square error form using eqn. 10 as

$$\varepsilon_{ms} = \int_{-\pi}^{\pi}\varepsilon^2(\phi)\,d\phi = \frac{1}{\pi}\sum_{n=T+1}^{\infty} b_{(nL)}^2 \quad (17)$$

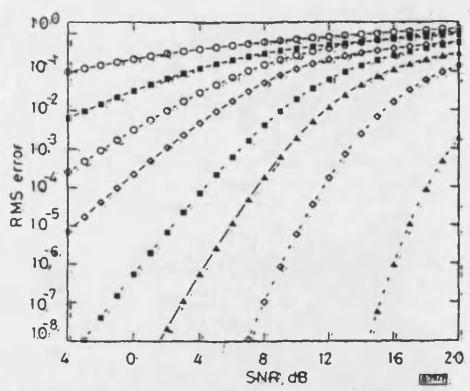


Fig. 1 RMS error in PDF against SNR for multilevel PSK with one and three term approximations

○ 2 PSK
■ 4 PSK
◆ 8 PSK
▲ 16 PSK
——— 1 term
- - - - 3 term

Fig. 1 shows the RMS error as a function of SNR for a one and three term approximation; it can be seen that as the number of terms is increased, the approximation becomes significantly more accurate.

Conclusions: This Letter has shown that the phase PDF of multilevel PSK may be expressed as a DC component and a sinusoid for low SNR, and a plot has been given for the RMS error against SNR. For an increased SNR range, the number of sinusoid terms may be increased. The scaling of these sinusoids may be determined using a simple recurrence property of Bessel functions.

Acknowledgments. The authors would like to thank Vodafone Plc. and SERC for the funding of this research.

9 May 1994
Electronics Letters Online No: 19940882

P. C. Sapiano, R. J. Holbeche and J. D. Martin (*School of Electronic & Electrical Engineering University of Bath, BA2 7AY, United Kingdom*)

References

1 WHALEN, A.D.: 'Detection of signals in noise' (Academic Press, London, 1971), Chap. 4
2 LEIB, H., and PASUPATHY, S.: 'The phase of a vector perturbed by Gaussian noise and differentially coherent detectors', *IEEE Trans.*, 1988, IT-34, (6), pp. 1491–1501
3 GRADSHTEYN, I.S., and RYZHIK, I.M.: 'Tables of integrals, series and products' (Academic Press, London, 1980), Chap. 1.32
4 ABRAMOWITZ, M., and STEGUAN, I.A.: 'Handbook of mathematical functions' (National Bureau of Standards, 1965), Chap. 10

Further results in the classification of PSK signals using the optimum method

P.C. Sapiano, J.D. Martin and R.J. Holbeche

Indexing terms: Phase shift keying, Statistics for communications

The automatic classification of PSK signals using the maximum likelihood approach has been described by Tang and Soliman (1991). The Letter develops the results further from a two-class case into a multiclass case by examining techniques for deriving the probability of misclassification. Results are given for a four-class case which may be used to compare with other methods.

Introduction: There have been a number of methods presented for classifying the number of levels in a PSK signal [1-3]. All of these methods are capable of distinguishing between two types of PSK signal, and [1, 3] show how a greater number of PSK types may be classified. The method in [1] is based on a maximum-likelihood approach and gives results for the error probability when BPSK and QPSK are compared in the classification process. The method in [3] uses statistical moments of the phase PDF and extends the results from CW to 8PSK.

This Letter describes a method for numerically evaluating the probability of misclassification using the algorithm in [1] when more than two PSK types are included in the classification process. Calculated and simulated results are given for the cases of CW through to 8PSK. The results enable comparisons between the methods of [1, 3] and other emerging methods.

Development: The classification procedure takes a series of phase samples and passes each consecutive sample through a set of functions related to each considered PSK type. The results are then summed against each considered PSK type, and the choice of classified signal is based on the maximum of these results. In the technique it is assumed that the carrier frequency is accurately known, the probability of occurrence of each class is identical, all the samples are i.i.d with zero mean and the noise is AGWN.

The PDF of phase ϕ of a carrier wave in the presence of AGWN is

$$f(\phi) = \frac{1}{2\pi} e^{-\rho} \left[1 + \sqrt{\rho\pi}\cos\phi e^{\rho\cos^2\phi}(1 + \mathrm{erf}[\sqrt{\rho}\cos\phi])\right] \quad (1)$$

and for L-level PSK signals the PDF is extended to

$$f_L(\phi) = \frac{1}{L}\sum_{k=0}^{L-1} f\left(\phi + \frac{2\pi(k+0.5)}{L} - \pi\right) \quad (2)$$

Each of the M incoming phase samples is tested against all possible PSK classes β_j, which is the number of PSK levels for the class. The classified modulation type is chosen from the maximum *a priori* probability which is expressed below using Bayes theorem and the assumption that all modulation types are equally likely:

$$\max[p(\beta_j|\phi(i))] = \max[p(\phi(i)|\beta_j)] \quad (3)$$

which may be written in log-likelihood form as

$$l_{\beta_j} = \sum_{i=1}^{M} \ln[p(\phi(i)|\beta_j)] \quad (4)$$

When M is large, the distribution of l_{β_j} tends towards a normal distribution by virtue of the central limit theorem, and the mean and variance of the distribution are characterised by

$$\mu_{\beta_j\alpha} = M m_{\beta_j\alpha}(1) \quad (5)$$

$$\sigma^2_{\beta_j\alpha} = M\{m_{\beta_j\alpha}(2) - [m_{\beta_j\alpha}(1)]^2\} \quad (6)$$

where $m_{\beta_j\alpha}(n)$ is the nth moment of l_{β_j} when α-level PSK is transmitted. It can be shown that

$$m_{\beta_j\alpha} = \int_0^{\frac{\pi}{\beta_j}} (\ln[f_{\beta_j}(u)])^n \tilde{f}_\alpha(u)du \quad (7)$$

where $\tilde{f}_\alpha(u)$ is the PDF of phase of the true PSK signal ($f_\alpha(u)$), aliased about $u = \pi/\beta_j$ and positive u. This is in order to take into account the fact that $f_\alpha(u)$ is not a unique function. The mean and variance may be evaluated using eqns. 5 and 6 through the numerical integration of eqn. 7.

Probability of error: The probability of correct classification when the signal is in the interval $x, x + \delta x$ with α-level PSK transmitted is given by $k_\alpha(x)\delta x$, where

$$k_\alpha(x) = \frac{1}{\sigma_{\alpha\alpha}\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_{\alpha\alpha})}{2\sigma^2_{\alpha\alpha}}\right] \prod_{\substack{j \\ \beta_j \neq \alpha}} \Phi\left(\frac{x-\mu_{\beta_j\alpha}}{\sigma_{\beta_j\alpha}}\right) \quad (8)$$

and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \quad (9)$$

The overall probability of correct classification given that α-level PSK is transmitted is given in the limit of $\delta x \to 0$:

$$p(\text{correct}|\alpha T.x) = \int_{-\infty}^{\infty} k_\alpha(x)dx \qquad (10)$$

As $\Phi(x)$ is a linearly increasing function ranging from 0 to 1, $k_\alpha(x)$ becomes insignificant within the limits $\mu_{\alpha\alpha} - 20\sigma_{\alpha\alpha} > x > \mu_{\alpha\alpha} + 20\sigma_{\alpha\alpha}$. Therefore eqn. 10 may be simplified to

$$p(\text{correct}|\alpha T.x) = \int_{\mu_{\alpha\alpha}-20\sigma_{\alpha\alpha}}^{\mu_{\alpha\alpha}+20\sigma_{\alpha\alpha}} k_\alpha(x)dx \qquad (11)$$

Finally the probability of error may be numerically evaluated through eqn. 11 and using

$$p(\text{error}|\alpha T.x) = 1 - p(\text{correct}|\alpha T.x) \qquad (12)$$

This method has been used to calculate the error probability for CW, BPSK, QPSK and 8PSK, which are given in Fig. 1*a* and *b*. The error probability was also deduced through simulation of 10000 trials, and these results are also shown in the Figures. The results for 8PSK are given separately in Fig. 1*b* as they lie close to the QPSK points. A similar effect was found in [1], where the BPSK and QPSK misclassification probabilities were also numerically similar.

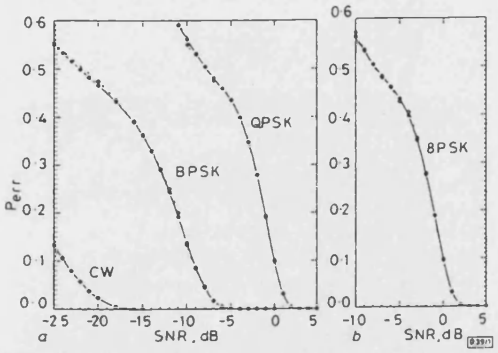


Fig. 1 *Misclassification probability against SNR for CW, BPSK, QPSK and 8PSK*

a CW, BPSK, QPSK
b 8PSK
—●— calculated results
– –◇– – simulated results

The evaluation of the PDF functions in [1] was through a Fourier series expansion. These coefficients may be evaluated efficiently using the method detailed in [4]. Other methods of evaluation may be obtained by using numerical approximations to the error function [5] and evaluating the function directly from eqn. 2. When the expressions applied to eqn. 2 are expanded and simplified, computational speed improvements may be obtained.

The results in Fig. 1*a* and *b* now enable a comparison between the techniques in [1, 3]. For the four-class case the optimum method [1] outperforms the method of statistical moments [3] in terms of error performance. A numerical comparison is drawn between the two techniques when [3] uses the 8th statistical moment. This is expressed as the difference in SNR required for 1% error probability, and is summarised in Table 1.

Table 1: SNR gain of optimum method over 8th statistical moment for 1% error probability and four-class case

	CW	BPSK	QPSK	8PSK
SNR gain [dB]	4.2	3.0	4.4	4.6

Conclusions: The optimum method for classifying multilevel PSK signals using the maximum-likelihood method has been examined. A technique for evaluating the probability of misclassification is derived for two classes in [1] and for more than two classes in this Letter. Results are produced which have been verified through computer simulation and have been used to compare with the method of statistical moments for a four-class case, where it is found that the optimum method has improved performance.

Acknowledgments: The authors would like to express thanks to Vodafone plc and the UK SERC for the funding of this research.

21 October 1994
Electronics Letters Online No: 19950047

P.C. Sapiano, J.D. Martin and R.J. Holbeche (*School of Electronics & Electrical Engineering, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom*)

References

1 TANG, Y., and SOLIMAN, S.S.: 'Optimum classifier for M-ary PSK signals'. IEEE Int. Conf. on Communications, 1991, pp. 1693–1697
2 POLYDOROS, A., and KIM, K.: 'On the detection and classification of quadrature digital modulations in broad-band noise', *IEEE Trans.*, 1990, **COM-38**, (8), pp. 1199–1211
3 SOLIMAN, S.S., and HSUE, S.Z.: 'Signal classification using statistical moments', *IEEE Trans.*, 1992, **COM-40**, (5), pp. 908–916
4 SAPIANO, P.C., HOLBECHE, R.J., and MARTIN, J.D.: 'Low SNR approximation to phase PDF for PSK signals', *Electron. Lett.*, 1994, **30**, (16), pp. 1279–1280
5 ABRAMOWITZ, M., and STEGUN, I.A.: 'Handbook of mechanical functions' (National Bureau of Standards, 1965), Chap. 7

CLASSIFICATION OF PSK SIGNALS USING THE DFT OF PHASE HISTOGRAM

P.C. Sapiano, J.D. Martin and R.J. Holbeche
School of Electronic Engineering, University of Bath,
Bath, BA2 7AY, England

ABSTRACT

A method is presented for classifying multi-level PSK signals in the presence of additive white Gaussian noise (AGWN). The technique is based on the Discrete Fourier Transform (DFT) of a phase histogram. The probability of correct classification is given and it is found that the technique performs well at low SNR. The benefits of this technique are that it is simple to implement and requires no prior knowledge of the SNR of the signal for the classification.

1. INTRODUCTION

The automatic classification of modulation type of a communications signal finds applications in the fields of Electronic Surveillance, spectrum management and signal interception where it is an important sorting parameter in a complicated problem. There are also applications in modulation diverse communication systems in which the system may receive a variety of modulation types. The initial trend was to treat modulation recognition as a non-deterministic pattern recognition problem [1] which works well at a high SNR, but is poor at low SNR. Consequently some research effort has been applied to deterministic forms of pattern recognition [2][3][4]. Some of these methods rely upon knowledge of the SNR of the signal for decision parameters, but this is difficult to obtain in a true signal environment. The method presented shows a technique for classifying PSK signals without using SNR information, and the performance is found to work well at low SNR.

Phase samples of the incoming signal are collected and placed in a histogram. After a sufficient number of these have been collected, the histogram is passed through a DFT in order to exploit the periodicity of the histogram. The DFT bin numbers correspond to the number of levels of each of the PSK types considered, and are converted into magnitude squared where the largest of these identifies the PSK type (figure 1).

The theoretical development given below shows that the number of bins should be a power of two to avoid the effects of spectral leakage, and a closed form expression for the probability of misclassification is derived. It is found that the number of histogram bins used in the process does not affect the classification procedure when aliasing effects are made insignificant. Finally the error probability for this method is compared to that of the method of statistical moments, and it is found to compare favourably.

2. THEORETICAL DEVELOPMENT

A time frame of the signal of interest is captured and digitised. This signal is then converted into an analytic form, the carrier is removed by complex mixing and the phase samples are extracted. It is assumed that the carrier frequency component is accurately known. The received signal is the sum of an ideal PSK signal and AGWN. The phase p.d.f. of multi-level PSK in AGWN may be developed from a carrier wave (CW) p.d.f. $f(\phi)$, which is described in Fourier series form as [5]:

$$f(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{n=1}^{\infty}b_n\cos(n\phi) \quad -\pi<\phi\le\pi \tag{1}$$

When M level PSK is considered, the p.d.f. $f_M(\phi)$ becomes :

$$f_M(\phi)=\frac{1}{M}\sum_{k=0}^{M-1}f\left(\phi+\frac{2\pi(k+0.5)}{M}-\pi\right) \tag{2}$$

The Fourier series form of this p.d.f. is given by [6]

$$f_M(\phi)=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^n b_{(nM)}\cos(nM\phi) \tag{3}$$

The Fourier series coefficients b_m are a function of the SNR ρ, which is given by [5] as :

$$b_m=\frac{\sqrt{\rho\pi}e^{-\frac{\rho}{2}}}{2}\left[I_{\frac{mM}{2}}\left(\frac{\rho}{2}\right)+I_{\frac{mM}{2}-1}\left(\frac{\rho}{2}\right)\right] \tag{4}$$

$I_v(x)$ is the modified Bessel function of order v, which is of an integer plus a half order. It is found that as m is increased b_m decreases, and as ρ is decreased the separation between harmonic magnitudes increases [6]. At low SNR the Fourier series is dominated by the mean and first harmonic. This final property will be used when evaluating the probability of false classification.

0-7803-2431-5/95 $4.00 © 1995 IEEE

2.1 Histogram Representation

The phase samples $\phi(n)$ are used to build up a phase histogram with N bins and L samples to approximate the p.d.f.. The following theory characterises the error between the true p.d.f. and the histogram approximation.

It is known [7] that the variance σ_i^2 between the true p.d.f. and the histogram estimate for a particular histogram bin i is :

$$\sigma_i^2 \approx \frac{1}{L\Delta} f(\phi_i) \qquad (5)$$

where L is the number of samples, Δ is the bin width, which is assumed to be small, and i is the histogram bin number. By virtue of the central limit theorem, the errors of all the bins will be normally distributed for large L. The mean variance of the error terms is expressed by :

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \qquad (6)$$

where $\hat{\sigma}^2$ is the noise variance of the N bin histogram. As the histogram has equally spaced points in the interval $[-\pi,\pi]$,

$\Delta = \frac{2\pi}{N}$. It can be shown that the noise variance is :

$$\hat{\sigma}^2 \approx \frac{N}{4\pi^2 L} \qquad (7)$$

2.2 Discrete Fourier Transform

The N points of the histogram are operated on by the discrete time Fourier transform to exploit spectral peaks corresponding to the harmonic terms. N is made to be a power of 2 in order that the harmonic terms will coincide with the frequency bins, thus avoiding the effects of spectral leakage. The bin number corresponds to the harmonic number (where the D.C. component is on bin 0), and from (1) it is seen that M level PSK will be characterised by a series of spectral lines on the bins which are a multiple of M.

It can be shown by the sampling theorem that harmonics of order $\frac{N}{2}$ and higher will be subjected to aliasing. In order to avoid the aliasing of the fundamental harmonics for any of the PSK schemes presented to the system for classification, N must be at least four times the highest symbol number.

The magnitude squared of the DFT is used as it is simple to calculate, and this is scaled by $\frac{4}{N^2}$ to produce D(k) which will provide magnitude squared values for the harmonic terms. When a harmonic component is not present in a bin that bin has only AGWN present and D(k) will be Rayleigh distributed [8] with variable y and p.d.f. :

$$p(y) = \pi^2 L \exp(-\pi^2 L y) \quad y>0 \qquad (8)$$

This result shows that the noise floor is independent of the number of histogram bins N and implies that N may be made large enough to remove any significant effects of aliasing without affecting the noise floor.

When a frequency bin is occupied by a harmonic signal with the histogram noise, the bin is distributed with a non-central Chi-squared distribution, with two degrees of freedom [9]. It can be shown that the distribution g(x) of a bin D(k) containing a harmonic of amplitude $\frac{b_m}{\pi}$ and the histogram noise is given by :

$$g(x) = \pi^2 L \exp\left(-L\left[b_m^2 + \pi^2 x\right]\right) I_0\left(2\pi b_m L\sqrt{x}\right) \quad x>0 \qquad (9)$$

Where $I_0(z)$ is the modified Bessel function of zero order. It should be noted that this expression is also independent of N.

2.3 Classification

The classification is achieved by finding the maximum DFT magnitude for the bins which are of interest, $D(\alpha_n)$ where α_n is the number of states in the n^{th} PSK signal. The classified signal is M-PSK where :

$$\alpha_M \in MAX\left[D(\alpha_n)\right] \qquad (10)$$

e.g. when 1,2,4 & 8 PSK are to be classified, bins 1,2,4 & 8 of D(k) are examined, and if bin 4 is the maximum then the signal is classified as 4 PSK.

2.4 Probability Of False Classification

Consider the bin containing the signal x with distribution g(x), and n noise bins which are i.i.d. with distribution p(y). The probability that the signal lies in the interval $x, x+\delta x$ is given by

$$g(x)\delta x \qquad (11)$$

The condition for correct classification is that the noise signals are less than x. The probability of correct classification in the interval is therefore :

$$g(x)[1-\Phi(x)]\delta x \qquad (12)$$

where

$$\Phi(x) = \int_x^{\infty} p(y)dy \qquad (13)$$

When all of these contributions are summed and in the limit of $\delta x \to 0$, the probability of correct classification is given by :

$$p_{corr} = \int_0^{\infty} g(x)[1-\Phi(x)]^n dx \qquad (14)$$

Which can be re-written as :

$$p_{corr} = \sum_{i=0}^{n} \frac{n!(-1)^i}{(n-i)!i!} \int_0^{\infty} g(x)\Phi^i(x)dx \qquad (15)$$

The probability of error is given by :

$$p_{err} = 1 - p_{corr} = \sum_{i=1}^{n} \frac{n!(-1)^{i+1}}{(n-i)!i!} \int_0^{\infty} g(x)\Phi^i(x)dx \qquad (16)$$

From (8), $\Phi(x)$ is given by :

$$\Phi(x) = \exp(-\pi^2 Lx) \qquad (17)$$

Using (9) and (17) and [10] it can be shown that

$$\int_0^{\infty} g(x)\Phi^i(x)dx = \frac{1}{i+1}\exp\left[-\frac{ib_m^2 L}{i+1}\right] \qquad (18)$$

Therefore the error probability is given by :

$$p_{err} = \sum_{i=1}^{n} \frac{n!(-1)^{i+1}}{(n-i)!(i+1)!}\exp\left[-\frac{ib_m^2 L}{i+1}\right] \qquad (19)$$

3. RESULTS

Plots of error probability against SNR are given in figure 2 for the case of CW, BPSK, QPSK and 8PSK being examined with a sample length of 1024 points. This is compared with simulated results, and it is found that the model is accurate for CW and BPSK, but QPSK deviates slightly, and 8PSK deviates further from the theory which indicates that the Gaussian assumption of the noise becomes less accurate. However the simulated results show that the error probability is better than that which the theoretical model suggests, and the two tend to converge at error probabilities less than 1% which are the main areas of interest.

This is compared with classification using the 8^{th} statistical moment [2] (figure 3) and at a 1% error probability. The first column of Table 1 shows the SNR gain of the new technique and it is found that the new technique is better in every case except 8 PSK. This is because the statistical moments technique assumes that a signal with a moment greater than that of 8PSK will not be present. The second column is the same comparison when 16 PSK is also included, and it is found that the new technique performs better in all cases.

Finally a comparison is drawn with the method of maximum likelihood classification [4] (figure 4) and it is found that the proposed method is close in error performance in all cases except that of 8PSK where the method is outperformed by 6.5 dB and lies close to the error performance of QPSK. It should be noted that the 'optimum' method requires a heavy computational overhead for each sample along with knowledge of the SNR of the signal.

4. CONCLUSIONS

A new method has been presented for the classification of multi-level PSK signals which requires no prior knowledge of the SNR of the signal unlike other deterministic methods. The algorithm is extremely simple and fast to implement requiring no complicated thresholding calculations. The error performance is found to be good and in certain cases it out-performs more complicated techniques. The technique works well at high and low SNR, and is proposed as an attractive method for the classification of PSK signals.

5. REFERENCES

[1] F. Jondral, "Automatic Classification Of High Frequency Signals," Signal Processing, vol.9, no.3, pp. 177-190, Oct. 1985.

[2] S.S. Soliman, S.Z. Hsue, "Signal Classification Using Statistical Moments," IEEE Trans. Communications, vol. 40, no. 5, pp. 908-916, May 1992.

[3] A. Polydoros, K.K. Kim, "On The Detection And Classification Of Quadrature Digital Modulations In Broad Band Noise," IEEE Trans. Communications, vol.38, no.8, pp. 1199-1211, Aug. 1990.

[4] Y. Tang, S.S. Soliman, "Optimum Classifier for M-Ary PSK Signals," IEEE International Conference on Communications, pp. 1693-97, 1991.

[5] N.M. Blachman, "Gaussian Noise-Part II: Distribution Of Phase Change Of Narrow-Band Noise Plus Sinusoid," IEEE Trans. Information Theory, vol. 34, no. 6, pp. 1401-1405, Nov. 1988.

[6] P.C. Sapiano, R.J. Holbeche, J.D. Martin, "A Low SNR Approximation To The Phase PDF For PSK Signals," IEE Electronics Letters, vol. 30, no. 16, pp.1279-1280, Aug. 1994.

[7] K.S. Shanmugan, A.M. Breipohl, "Random Signals, Detection, Estimation And Data Analysis," Wiley, ch. 8, 1988.

[8] R. Shiavi, "Introduction To Applied Statistical Analysis," Aksen Associates Incorporated Publishers, ch. 7, 1991.

[9] A.D. Whalen, "Detection Of Signals In Noise," Academic Press, New York., 1971.

[10] I. S. Gradshteyn and I.M. Ryzhik, "Table Of Integrals And Series Products," Academic Press, London, ch.6.631.4, 1980.

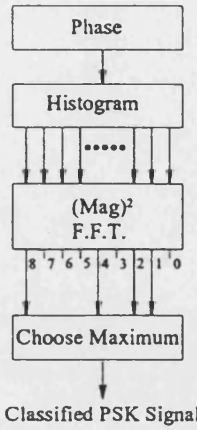


Figure 1 : Algorithmic description

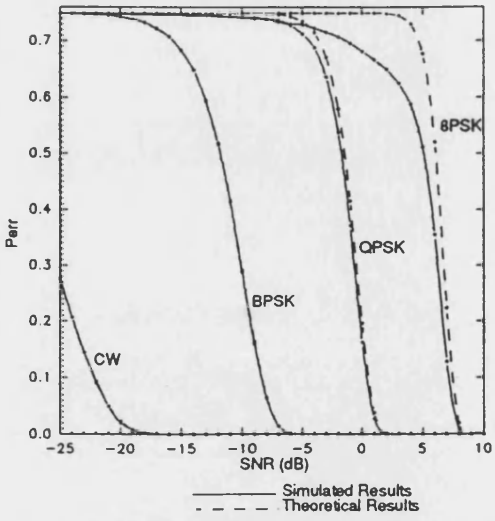


Figure 2 : Plots of classification error probability against SNR For CW, BPSK, QPSK, 8PSK for the DFT of phase histogram method

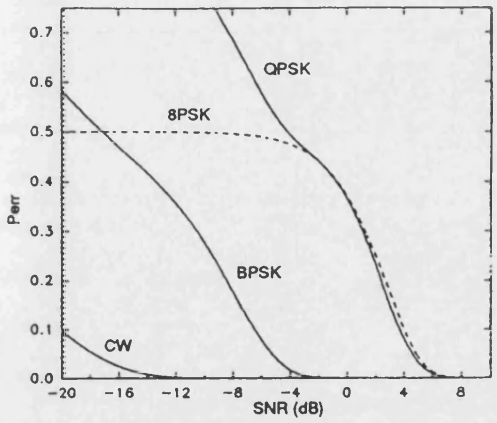


Figure 3 : Plot of error probability v SNR for the statistical moment classifier

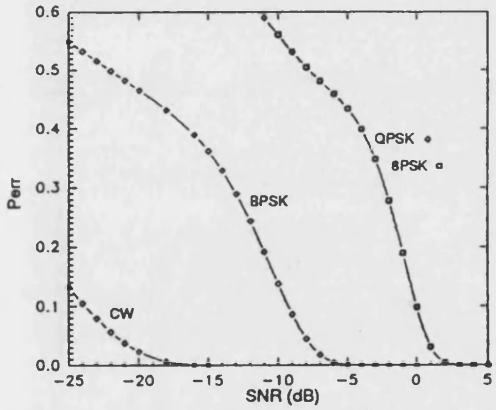


Figure 4 : Plot of error probability v SNR for the optimum classifier

	CW-8PSK	CW-16PSK
CW	4.3	3.9
BPSK	3	2.7
QPSK	4.3	4.2
8PSK	-2.5	Large
16PSK		Large

Table 1 : SNR gain (dB) of presented method, compared to 8^{th} statistical moment when the error probability is 1%

IDENTIFICATION OF PSK SIGNALS

P C Sapiano, J D Martin,

University of Bath, UK

ABSTRACT

PSK signals may be identified using decision theoretic techniques. This paper compares the performance of the optimum, the statistical moments, the DFT and the maximum likelihood DFT classifiers.

The robustness of each classifier is examined for the effects of symbol imbalance due to a finite signal time frame, error in the SNR estimate, channel filtering and phase error. Simulation results are presented in terms of the SNR at which 1% misclassification probability occurs, in order to provide a comparison between the techniques.

1. INTRODUCTION

Automatic modulation recognition finds applications in military and civil surveillance, spectrum management, interference identification, radio direction finding and modulation diverse communication systems [1][2].

Using decision theoretic techniques for automatic modulation recognition requires parametric modelling of the signal characteristics. It is useful in such cases to adopt a divide and conquer form of pattern classification where an analysis of the cyclostationary characteristics of the signal may identify the signal as a PSK form for example, indicating that the PSK classification algorithm should be implemented.

The classification of PSK signals using decision theoretic techniques determines the number of levels of the PSK signal [3],[4],[5] and [6]. In all of the algorithms it is assumed that the system has achieved carrier synchronisation, as a moderate frequency offset will obscure the phase signal. If perfect phase unwrapping is achievable, then the carrier component may be easily identified using linear regression on the mean phase ramp component. However phase unwrapping becomes extremely difficult to achieve under noisy conditions, and becomes even more difficult when there are phase transitions due to PSK signalling. This presents a conceptually difficult problem to solve, particularly when the received signal may not solely be a PSK form of signal.

In the work presented the signal is of a PSK form and it is assumed that either CW, BPSK, QPSK or 8PSK are presented for classification, and performance is evaluated through computer simulation with signal segments of 1024 points. Various PSK classification algorithms are evaluated under certain effects which have not previously been introduced in the literature. These effects will degrade the system in a broad manner, and the analysis is aimed to give an understanding of the robustness of the systems in a practical implementation.

The Optimum classifier [3], uses a likelihood function applied to the phase samples, and is based upon the pdf characteristics of the phase signal. The Statistical moments classifier [4] uses statistical moments to transform the signal into feature space, and is followed by a threshold process to determine the PSK type. In this work the 8th statistical moment is examined. The DFT classifier [5] is the simplest and fastest of the classifiers and analyses the pdf of phase through the DFT of the phase histogram. Finally the Maximum likelihood DFT classifier [6] uses a maximum likelihood function applied to analysis in [5].

The four effects which are considered are an asymmetry in symbol probabilities due to finite signal length (2.1), an error in the SNR estimate (2.2), band limiting on the signal and noise (2.3) and an error in the zero phase extraction (2.4).

2. DEVELOPMENT

2.1 Random Symbols

In the algorithms it is assumed that the symbol probabilities are equiprobable, where as in fact a finite time frame will cause the occurrence of each symbol to be unbalanced and consequently the effective probability of each symbol observed within the time-frame is not equiprobable.

The signal has been captured within a finite time-frame where N symbol transitions will occur. Assuming that the occurrence of each symbol level is independent and equiprobable, the symbols will be Binomially distributed with mean :

$$\mu_s = \frac{1}{M} \qquad (1)$$

and variance :

$$\sigma_s^2 = \frac{1}{N}\left(\frac{M-1}{M^2}\right) \qquad (2)$$

Radio Receivers and Associated Systems, 26-28 September 1995, Conference Publication No 415 © IEE 1995

As N is in general >20, this tends towards a normal distribution of the same mean and variance.

With the DFT forms of classifier, the harmonic power is distributed across the pdf frequency band which reduces the fundamental harmonic and increases the power in the off-harmonic bins, which along with the histogram noise power, contribute to the misclassification probability.

With the statistical moment method the statistical moments will become altered under such effects, thus moving the mean value towards one of the thresholds, and consequently increasing the error probability.

The optimum classifier is expected to suffer the least under the effect of offset symbol probabilities as the symmetry in the likelihood estimation functions acts to reduce the offset. e.g. in the case of 8PSK, when placed through the functions of 8PSK and QPSK, there is no effect due to offset symbol probabilities. When compared to BPSK, the effect is that the effective variance is a quarter of the symbol variance, as half of the symbols are passed through an identical function and with CW the effective variance is a half of the individual symbol variance. Therefore the effect of symbol imbalance is reduced by the optimum classifier.

For the purposes of evaluation a 1024 point time frame and the sampling rate to symbol rate ratio is varied up to a factor of 32. As an example of the symbol variance, a ratio of 4 produces 256 symbol transitions, and the means and 95% confidence limit ranges are described below :

CW	BPSK	QPSK	8PSK
0	0.5±0.061	0.25±0.053	0.125±0.041

The results for this are shown in figures 1a,b,c for sampling frequency ratios up to 32. Note that the optimum classifier is not included. This is because there was no significant change in performance across the tests. The maximum likelihood DFT classifier appeared to be robust, suffering a maximum of 2dB loss in performance. The DFT classifier performed worse, with up to 6dB worth of degradation, and the statistical moments classifier performed the worst, where the QPSK and 8PSK would converge to an error greater than 1% at high SNR.

2.2 SNR offset

The classifiers in [3],[4] and [6] assume that the SNR is accurately known, but in practice only a crude estimate of the SNR may be made, and the effects of this error in the estimate are examined.

As the DFT classifier [5] does not use SNR information, its characteristics are fixed for this test. The penalty that this algorithm pays for the lack of knowledge is an increased error probability for the highest PSK symbol, therefore the SNR error at which the DFT classifier outperforms the other classifiers is of interest.

This may be calculated for the optimum method using the technique in [7] by placing the estimated SNR in $f_{\beta_i}(.)$ of equation (7), and in the statistical moments classifier by placing the estimated SNR into the threshold calculations.

The plots are presented in figures 2 a,b,c as the estimated offset against the SNR at which a 1% error probability occurs for a particular PSK type. This is plotted for an offset range of -5dB to 5dB. Overplotted on the graphs are the characteristics of the DFT classifier, which remain constant against SNR offset.

The general trend is that the performance for 8PSK improves with an increased bias on the estimate, whereas all the other types perform worse. The statistical moments classifier is the most sensitive to this parameter. The maximum likelihood DFT classifier performs the best, but it can be seen that all are sensitive to this offset.

2.3 Channel Filtering

The algorithms used for classification have assumed that the noise is white and the symbols are stepped. In a real system, this will not be the case due to bandlimiting and non-linearities.

An filter has been implemented as a raised cosine filter with impulse response {0.25,0.5,0.25}, which is simple to implement and applies significant filtering on the signal. The overall noise power is reduced by 4.26 dB with such a filter. The symbol rate of the filtered signal is a quarter of the sampling frequency. In order to isolate the effects of unbalanced symbols, a PRBS of length 256 is used to determine each of the symbols such that the resultant signal has equally balanced symbol probabilities.

Band-limiting will cause the noise samples to no longer be independent, but will instead be correlated by a factor determined by the impulse response of the channel at the sampling intervals. The edges of the PSK transitions will not be square, but will follow a path related to the previous symbols and the channel impulse response and adjacent samples will no longer be independent. However the effects of bandlimiting will cause the overall noise power will be reduced, which is in the favour of the classification.

These effects will cause the pdf to be altered and the true maximum likelihood functions to deviate from those implemented.

In order to examine the effects of the filtering on the trajectory of the PSK signal in the absence of noise, the PSK signal may be considered as a narrowband process. The response of the channel g(t) to a step going from amplitude A_1 to A_2 is defined by h(t) in :

$$h(t) = A_2 \int_{-\infty}^{t} g(y)dy + A_1 \int_{-\infty}^{-t} g(y)dy = \frac{A_1 + A_2}{2} + p(t)\frac{A_1 - A_2}{2} \quad (3)$$

Where p(t) is a function extracted from h(t). It may be shown that the phase $\phi(t)$ follows a trajectory due to a step input given by :

$$\phi(t) = \phi_0 + \tan^{-1}\left(\frac{\sin(\Delta)[1+p(t)]}{1-p(t)+\cos(\Delta)[1+p(t)]}\right) \quad (4)$$

where Δ is the phase shift between adjacent symbols and ϕ_0 is the initial phase. In the particular example with a sampling frequency to symbol rate ratio of 4, no intersymbol interference is introduced and (4) may be used to determine the trajectory for each edge.

The simulation used the SNR prior to filtering. The results are given in table 1, where a general improvement in performance is observed, which may be attributed to the reduction in noise. From the results it may be deduced that parametric distortion is not present to a significant degree.

2.4 Phase Offset

The optimum [3] and statistical moments [4] classifiers assume that both the carrier frequency and phase are known. The phase may be recovered by removing the mean from the data set through post processing. This is however susceptible to error, which is attributed to two factors in an ideal system. The first is due to the finite number of samples used and the second error factor is related to the finite number of symbols as in section (2.1).

Each symbol is modelled by a mean phase perturbed by noise, and the pdf about the mean is that of the phase of a sinusoid in noise [8]. The variance of this is given in [8] by :

$$\sigma_n^2 = \frac{\pi^2}{3} + 4\sum_{i=1}^{\infty}\frac{(-1)^i}{i^2}b_i \quad (5)$$

Where b_i is the Fourier series coefficient for the pdf of phase of a sinusoid in noise, described in [8]. As each sample is i.i.d. and the number of samples L is large, the central limit theorem may be applied, where the overall mean estimate is the true mean with a normally distributed error of variance given by :

$$\sigma_{nL}^2 = \frac{\sigma_n^2}{L} \quad (6)$$

The second error due to the imbalance in symbol probabilities alters what was termed as the true mean. It is seen in section 2.1 that the mean probabilities of occurrence may be modelled as a binomial distribution. This may be taken to a limit of a normal distribution, and assuming that these are i.i.d., where the true mean is normally distributed about the desirable mean with variance given by :

$$\sigma_{sM}^2 = \frac{1}{N}\left(\frac{M-1}{M^3}\right) \quad (7)$$

where M is the PSK number. The overall error when the estimated mean is subtracted from the signal is given by the sum of the two individual errors and is again normally distributed, mean zero and variance given by :

$$\sigma_{off}^2 = \sigma_{nL}^2 + \sigma_{sM}^2 \quad (8)$$

Finally the standard deviation in radians :

$$\sigma_{off} = \frac{1}{\sqrt{L}}\left[\beta\left(\frac{M-1}{M^3}\right) + \frac{\pi^2}{3} + 4\sum_{i=1}^{\infty}\frac{(-1)^i}{i^2}b_i\right]^{\frac{1}{2}} \quad (9)$$

where β is the sampling rate to bit rate ratio. Using β=4, a 95% confidence limit for the phase error ranges between 1.3° to 6.8° for BPSK and 0.4° to 3.3° for 8PSK.

The results for phase shifts of 5° and 10° are given in table 2, and it can be seen that QPSK suffers the most under the measured conditions whereas the others remain relatively undisturbed. The Optimum classifier appears to be less susceptible than the statistical moments classifier.

3. CONCLUSIONS

The classifiers are robust to the effects of symbol probability imbalance for a typical case, but as the imbalance is increased, the statistical moments classifier has poor performance, the optimum classifier has little detectable performance loss, the maximum likelihood DFT classifier loses up to 2dB in performance and the DFT classifier looses up to 6dB.

The classifiers appear to be sensitive to an error in the SNR estimate, and in general an estimate biased below the true SNR performs better than one biased above.

With the filtering of the PSK symbols, there has been little loss in performance, and there is in general a gain in performance due to the reduction in the noise. The classifiers appear to be robust to the parametric distortion introduced.

Errors in the zero phase estimate appear to have little effect on performance for the ranges of error due to statistical averages, but for larger phase errors the classification performance will degrade significantly, and is prominent in the case of the statistical moments classifier.

These conclusions indicate that the classifiers are robust enough to be introduced into a practical system, but care should be taken when implementing the statistical moments classifier which is the most sensitive to the parametric variations.

4. ACKNOWLEDGEMENTS

The Authors would like to express thanks to EPSRC and Vodafone plc for the funding of this research.

REFERENCES

[1] Jondral F, 1985, "Automatic Classification Of High Frequency Signals", Signal Processing, Vol.9 No.3, 177-190.

[2] Hipp JE, 1986, "Modulation Classification Based On Statistical Moments", MILCOM '86, 20.2.1 - 20.2.6.

[3] Yang Y, Soliman SS, 1991, "Optimum Classifier For M-ary PSK Signals", ICC 91, 52.3.1-52.3.5.

[4] Soliman SS, Hsue SZ, 1992, "Signal Classification Using Statistical Moments", IEEE trans. Comms., Vol. 40, No. 5, 908-916.

[5] Sapiano PC, Martin JD, Holbeche RJ, 1995, "Classification of PSK Signals using the DFT of Phase Histogram", ICASSP-95, Vol. 3, 1868-1871.

[6] Sapiano PC, Martin JD, "Maximum Likelihood PSK classification using the DFT of phase histogram", Accepted for publication, IEEE GLOBECOM'95.

[7] Sapiano PC, Martin JD, Holbeche JD, 1995, "Further Results In The Classification Of PSK Signals Using The Optimum Method", IEE Electronics Letters, Vol. 31, No. 1, 19-20.

[8] Yang Y, Soliman SS, 1991, "Statistical Moments Based Classifier for MPSK Signals", GLOBECOM '91, 2.7.1-2.7.5

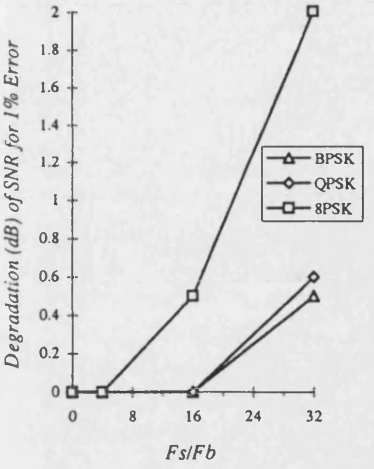


Figure 1a : Degradation of the 1% error probability v sampling rate to bit rate ratio for the Maximim likelihood DFT classifier

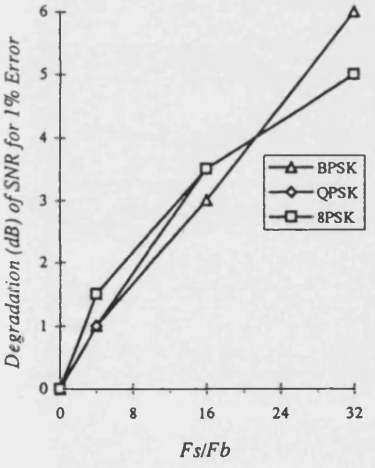


Figure 1b : Degradation of the 1% error probability v sampling rate to bit rate ratio for the DFT classifier

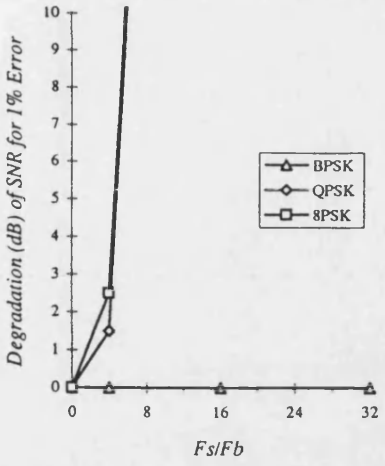


Figure 1c : Degradation of the 1% error probability v sampling rate to bit rate ratio for the statistical moments classifier

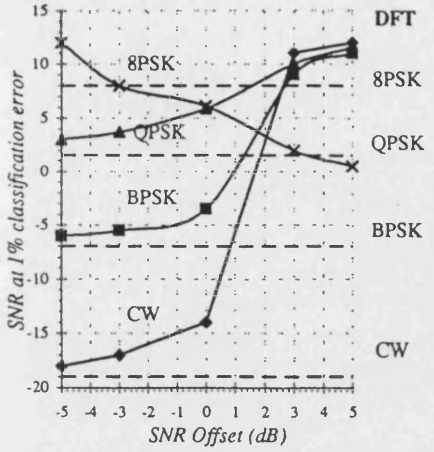


Figure 2b, Performance with SNR offset, Statistical Moments Classifier

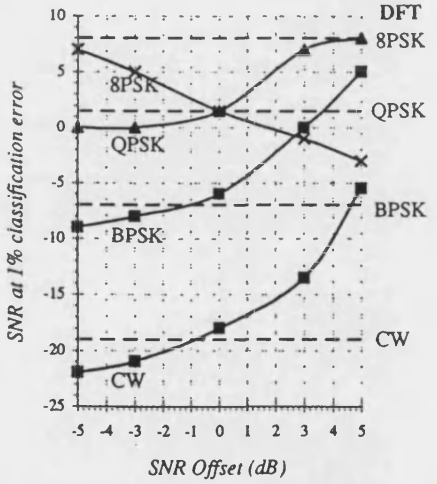


Figure 2a, Performance with SNR offset, Optimum Classifier

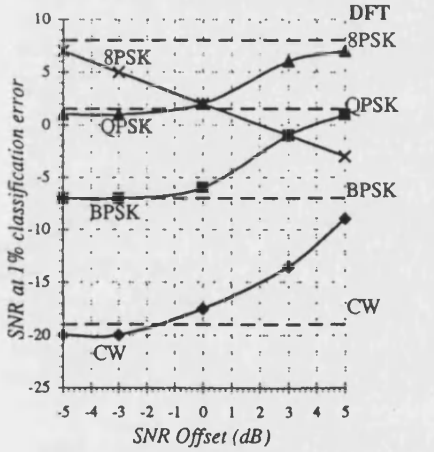


Figure 2c, Performance with SNR offset, Max. Likelihood DFT Classifier

	CW	BPSK	QPSK	8PSK
Optimum	3	4	4	-0.5
DFT	2	2	2.5	2
Statistical	5.5	4	4	0
Opt DFT	4	4	4	-0.5

Table 1, SNR improvement with raised cosine filtering

	CW	BPSK	QPSK	8PSK	Offset
Optimum	0	0	0	0	5 °
	0	0.5	2	0	10 °
Statistical	-0.5	-0.5	1	0	5 °
moments	1	1	7	-0.5	10 °

Table 1, SNR degradation due to a phase offset

MAXIMUM LIKELIHOOD PSK CLASSIFICATION USING THE DFT OF PHASE HISTOGRAM

P.C. Sapiano, J.D. Martin
University of Bath,Claverton Down, Bath, U.K. BA2 7AY

ABSTRACT

A method is presented for the classification of multilevel PSK signals which uses a maximum likelihood function on the DFT of phase histogram. An expression for this likelihood function is derived, which results in a simple function involving modified Bessel functions. The method has been evaluated for the classification of CW-8PSK and BPSK/QPSK and the error performance is comparable to or better than other methods used for comparison. However the method performs well in terms of computational complexity, which makes it attractive for the classification of PSK signals.

I. INTRODUCTION

The automatic classification of the modulation type of communications signals finds applications in military and civil surveillance, interference identification, radio direction finding and modulation diverse mobile radio systems. It is an area of research which has picked up some attention over the past decade and has seen a variety of techniques adopted. One area which has received some considerable attention is the problem of PSK classification [1][2][3][4], where the number of symbol levels of the PSK signal are classified. In such a problem it is desirable to employ deterministic pattern recognition principles for classification, as this results in improved classification performance at low SNR. For a more detailed discussion on the contributions to automatic modulation recognition, the reader is referred to [1] and [2].

The motivation of this paper is to introduce a new method for PSK classification which is based on a maximum likelihood classifier. In [1] the maximum likelihood classifier was based upon the phase samples directly. In this paper a further process is involved where the phase samples are placed into a histogram and the DFT of this histogram is used for the maximum likelihood estimation. It is found that the two methods have similar error performance, but the new method has significant gain in terms of computational complexity.

The technique is applied by collecting phase samples of the incoming signal and placing them into a histogram (figure 1), and a DFT is then applied to exploit the periodic structure of the phase pdf. In [3] the maximum DFT output indicates the classified PSK type, but in this further advance in the technique a likelihood function is applied before the selection process.

This paper is divided into the following format : Section II covers the theoretical development of the classifier where the pdf of phase for a multilevel PSK signal in AGWN is analysed in section II.A. The characteristics of the phase signal when placed into a histogram is examined in section II.B and the pdf for the output of the DFT function is derived in II.C. This then forms the basis for the maximum likelihood function which is derived in section II.D. The analytical evaluation of error probability is examined in section II.E. The results are discussed in section III, where the error performance is compared to various methods and finally the complexity of the algorithm is compared with that of [1].

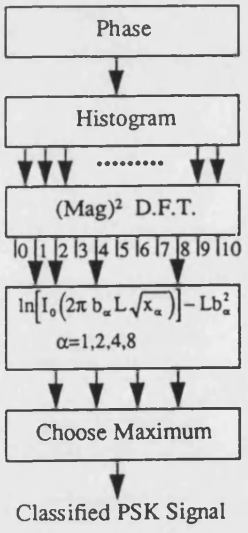


Figure 1 : Algorithmic description of the classifier

0-7803-2509-5/95 US$4.00 © 1995 IEEE

II. THEORETICAL ANALYSIS

The received signal $r(t)$ may be considered as the sum of an ideal PSK signal $s(t)$ and AGWN $n(t)$ with variance σ^2 as :

$$r(t) = s(t) + n(t) \tag{2.1}$$

The PSK signal is received with envelope A, carrier frequency f_c and a phase $\theta(t)$.

$$s(t) = A\cos(2\pi f_c t + \theta(t)) \tag{2.2}$$

The signal to noise ratio ρ may be expressed as :

$$\rho = \frac{A^2}{2\sigma^2} \tag{2.3}$$

$r(t)$ may be considered as a narrow band process [5] with envelope $z(t)$ and phase $\phi(t)$.

$$r(t) = z(t)\cos(2\pi f_c t + \phi(t)) \tag{2.4}$$

The received signal is sampled with sampling interval T_s to produce a sampled signal $r(n)$, and an analytic signal $\hat{r}(n)$ may be formed using Hilbert transform techniques [5], where :

$$\hat{r}(n) = z(n)\exp\left[j(2\pi f_c n T_s + \phi(n))\right] \tag{2.5}$$

Like all phase-based classifiers, it is assumed that f_c is accurately known, so the carrier frequency may be removed from the signal, to give $R(n)$ where

$$R(n) = z(n)\exp\left[j\phi(n)\right] \tag{2.6}$$

The angle of $R(n)$ is used to produce the sampled estimate of phase $\phi(n)$.

A. Phase Probability Density Function

The phase pdf of multilevel PSK in AGWN $f(\phi)$ may be developed from a carrier wave (CW) pdf, which has been described in [5] by

$$f(\phi) = \frac{1}{2\pi}e^{-\rho} + \frac{e^{-\rho}}{2}\sqrt{\frac{\rho}{\pi}}\cos(\phi)\exp[\rho\cos^2(\phi)]\left(1 + erf\left[\sqrt{\rho}\cos(\phi)\right]\right)$$
$$-\pi < \phi \le \pi \tag{2.7}$$

This may be represented as an even term Fourier series [6] as :

$$f(\phi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} b_n \cos(n\phi) \qquad -\pi < \phi \le \pi \tag{2.8}$$

It is assumed that each symbol has equiprobable occurrence. Therefore the pdf of M level PSK $f_M(\phi)$ may be described by :

$$f_M(\phi) = \frac{1}{M}\sum_{k=0}^{M-1} f\left(\phi + \frac{2\pi(k+0.5)}{M} - \pi\right) \tag{2.9}$$

The Fourier series representation is given in [7] as :

$$f_M(\phi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^n b_{(nM)}\cos(nM\phi) \tag{2.10}$$

The Fourier series coefficients b_m are a function of the SNR ρ, which are given in [6] as :

$$b_m = \frac{\sqrt{\rho\pi}e^{-\frac{\rho}{2}}}{2}\left[I_{\frac{m+1}{2}}\left(\frac{\rho}{2}\right) + I_{\frac{m-1}{2}}\left(\frac{\rho}{2}\right)\right] \tag{2.11}$$

$I_v(x)$ is the modified Bessel function of order v. An efficient method for generating these Bessel functions for the cases of multilevel PSK is described in [7] which uses reverse recursion techniques applied to spherical Bessel functions.

It is shown in [7] that as m is increased b_m decreases, and as ρ is decreased the separation between harmonic magnitudes increases. At low SNR the pdf may be represented by the mean and first harmonic term. This fact is used to derive the maximum likelihood classifier.

B. Histogram Probability Density Function

Each of L phase samples is placed into one of N histogram bins. This histogram is an approximation to the true pdf of phase for the incoming signal, and using certain assumptions the error between the histogram approximation and the true pdf may be characterised as AGWN [3] with variance $\hat{\sigma}^2$ given by :

$$\hat{\sigma}^2 \approx \frac{N}{4\pi^2 L} \tag{2.12}$$

C. DFT Bin Probability Density Function

The spectral components of the phase pdf Fourier series are revealed through a DFT on the histogram. N is set to be a power of two in order that the harmonic terms will coincide with the frequency bins, thus avoiding spectral leakage. The bin number of the DFT corresponds to the harmonic number of the phase pdf (where the D.C. component is on bin 0), and from (2.10) it is seen that M level PSK will be described by a series of spectral lines on the bins which are a multiple of M.

Harmonic terms of order $\frac{N}{2}$ and higher will be subjected to aliasing. The first harmonic of each scheme is used in the evaluation of the likelihood function, and aliasing is avoided by placing N at least four times the highest symbol number.

As the process requires the detection of the harmonic components, the magnitude squared of the DFT is used. The DFT bin which corresponds to the number of symbol levels of the transmitted PSK signal will contain the fundamental harmonic component, and it is assumed that all other bins contain no significant contribution from the

Fourier harmonics. This assumption improves at lower SNR, where it is critical. The DFT series is normalised about $\frac{4}{N^2}$ such that the resulting bin will equate to the magnitude squared of the Fourier series harmonic. This normalised DFT bin will be denoted by 'the DFT bin' hereafter.

The pdf of the mth DFT bin with the Fourier series harmonic may be represented by $g_m(x)$, which is a non-central chi-squared distribution with two degrees of freedom [3] as :

$$g_m(x) = \pi^2 L \exp\left(-L\left[b_m^2 + \pi^2 x\right]\right) I_0\left(2\pi b_m L\sqrt{x}\right) \quad x>0 \tag{2.13}$$

The pdf $p(x)$ of the DFT bins without a harmonic component are Rayleigh distributed and are given by [3] :

$$p(x) = \pi^2 L \exp\left(-\pi^2 L x\right) \quad x>0 \tag{2.14}$$

D. Maximum Likelihood Classifier

In the DFT classifier [3] the classification was based on the maximum of a set of DFT bins, each of which corresponded to a PSK type. This method required no knowledge of the SNR of the signal, but is sub-optimal in the case when the SNR information is known. The technique presented here assumes that the SNR information is available, and uses a maximum likelihood technique to classify the PSK type.

The probability that 2^α level PSK is transmitted given the DFT data is denoted by :

$$p\left(2^\alpha \text{ PSK} \mid x_0, x_1, \ldots x_\alpha, \ldots x_{max}\right) \tag{2.15}$$

Where x_n is the random variable representing the nth DFT bin corresponding to 2^n level PSK. Each PSK type is tested in turn using (2.15) and the type with the maximum result is chosen as the classified PSK type, i.e. :

$$\text{MAX}\left[p\left(2^\alpha \text{ PSK} \mid x_0, x_1, \ldots x_\alpha, \ldots x_{max}\right)\right]$$
$$\alpha \in \text{ all PSK types} \tag{2.16}$$

As all modulation types are considered equally likely, Bayes theorem can be applied to express the classified signal as :

$$\text{MAX}\left[p\left(x_0, x_1, \ldots x_\alpha, \ldots x_{max} \mid 2^\alpha \text{ PSK}\right)\right]$$
$$\alpha \in \text{ all PSK types,} \tag{2.17}$$

As each of the DFT bins have independent signals, the expression in (2.17) is evaluated as :

$$p\left(x_0, x_1, \ldots x_\alpha, \ldots x_{max} \mid 2^\alpha \text{ PSK}\right) = g_\alpha(x_\alpha) \prod_{\substack{i=0 \\ i\neq\alpha}}^{max} p(x_\alpha) \tag{2.18}$$

Therefore using (2.13) and (2.14) :

$$p\left(x_0, x_1, \ldots x_\alpha, \ldots x_{max} \mid 2^\alpha \text{ PSK}\right)$$
$$= \pi^2 L \exp\left(-L\left[b_\alpha^2 + \pi^2 x_\alpha\right]\right) \cdot I_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right) \prod_{\substack{i=0 \\ i\neq\alpha}}^{max} \pi^2 L \exp\left(-\pi^2 L x_i\right)$$
$$= \exp\left(-L b_\alpha^2\right) I_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right) \prod_{i=0}^{max} \pi^2 L \exp\left(-\pi^2 L x_i\right) \tag{2.19}$$

The terms independent of α may be eliminated, and the classified PSK signal can be represented through :

$$\text{MAX}\left[\exp\left(-L b_\alpha^2\right) I_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right), \alpha \in \text{all PSK types}\right] \tag{2.20}$$

Evaluating the Bessel function with large index can lead to mathematical overflow. This may be avoided by using the logarithm of the likelihood function (Which is a monotonically increasing function), and applying asymptotic expansions for the Bessel function. For large values, the Bessel function may be represented as [8] :

$$I_0(x) \approx \frac{1}{\sqrt{2\pi x}} \exp(x) \tag{2.21}$$

Therefore, the likelihood function ℓ_α may be represented by :

$$\ell_\alpha = \ln\left[I_0\left(2\pi b_\alpha L\sqrt{x_\alpha}\right)\right] - L b_\alpha^2 \quad x<100 \tag{2.22}$$

Which may be evaluated using approximations in [8]. For large argument the likelihood function is simply :

$$\ell_\alpha = L b_\alpha \left(2\pi\sqrt{x_\alpha} - b_\alpha\right) - \tfrac{1}{2}\ln\left[4\pi^2 b_\alpha L\sqrt{x_\alpha}\right] \quad x>100 \tag{2.23}$$

E. Misclassification Probability

The evaluation of the error probability involves the following integral :

$$P_{err} = 1 - \int_0^\infty r(x) \prod_j \left\{\int_0^x f_j(y)dy\right\} dx \tag{2.24}$$

Where $r(x)$ is the pdf of the output of the likelihood function for the DFT bin α, corresponding to the transmitted signal. $f_j(x)$ is the pdf of the likelihood function of DFT bin j for bins without signal presence. This evaluates into an integral of the following form :

$$\int \exp(-\nu x) \prod_j \left\{1 - \exp\left(\lambda\, I_0^{-1}\left[\beta_j\, I_0\left(\gamma\sqrt{x}\right)\right]\right)\right\} dx \tag{2.25}$$

Where β, λ, γ and ν are constants for the purpose of the evaluation. This integral is unfortunately extremely difficult to evaluate even using numerical methods, therefore the error performance has been evaluated

through computer simulation, the results of which are shown in figures 2 and 3.

III. RESULTS

Figures 2 and 3 show plots of error performance against SNR which have resulted from simulation trials. Figure 2 shows the case when L=1024 and CW, BPSK, QPSK and 8PSK are classified. Figure 3 shows the case where just BPSK and QPSK are classified. From these plots it can be seen that the error performance is good at low SNR.

A comparison between the various techniques in terms of error performance for the case when CW, BPSK, QPSK and 8 PSK are transmitted is given in table 1. This comparison is based upon the SNR at which there is a 1% misclassification probability. For a more detailed comparison, the reader is referred to the various texts [1][2][3] for the full plots against SNR.

The methods compared are the optimum method [1], with the results in question being published in [9], the method of statistical moments [2], where the 8^{th} statistical moment is used, and the DFT method [3].

It can be seen that the presented method performs better than the method of statistical moments in all cases, and the general performance is similar to that of the DFT method for CW-QPSK, but significantly better for the case of 8PSK. However, it should be noted that the DFT method does not require prior knowledge of the SNR of the signal unlike all the other methods considered.

Finally the performance is seen to be similar to that of the optimum classifier. This is also true of the case when BPSK and QPSK are used in the classification (table 2). Therefore the method looks to be potentially useful in terms of classification probability for PSK classification.

As the 'Optimum Classifier' and the new method have comparable error performance, a comparison on the algorithmic complexity is now given. The computation is divided into calculations which are required on a per sample basis and those which are required once off.

The new classifier requires only the phase samples to be placed into a histogram on a per sample basis, whereas the 'Optimum' classifier requires the evaluation of forty cosine terms and a logarithm for each class. On the once off calculations, both techniques require the generation of the Fourier series coefficients b_m, but the new method has the additional overhead of the evaluation of a DFT bin, square root, log function and on occasion a zero order Bessel function for each PSK class. When figures such as 1024 samples are used, the presented method has significant speed improvement due to the per sample speed improvement.

IV. CONCLUSIONS

A method has been presented which deals with the classification of multilevel PSK signals which is based upon the maximum likelihood classification of the DFT of phase histogram. The likelihood function has been derived in this paper and results of the algorithm are presented. It has been found that the algorithm performs well at low SNR, and has classification performance similar to that of the 'Optimum Classifier'. For a typical number of phase samples there is a significant improvement in computational complexity with the new technique over the 'Optimum Classifier'. Therefore the method is proposed as an attractive method for the classification of multilevel PSK signals.

ACKNOWLEDGEMENTS

The Authors would like to thank SERC and Vodafone plc for the funding of this research and Dr R.J. Holbeche for his assistance.

REFERENCES

[1] Y. Yang, S.S. Soliman, "Optimum Classifier For M-ary PSK Signals", Proc. IEEE International Conference On Communications, pp. 52.3.1-52.3.5, 1991.

[2] S.S. Soliman, S.Z. Hsue, "Signal Classification Using Statistical Moments", IEEE trans. Communications, Vol. 40, No. 5, pp. 908-916, May 1992.

[3] P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Classification of PSK Signals using the DFT of Phase Histogram", ICASSP-95, Vol. 3, pp. 1868-1871, May 1995.

[4] A. Polydoros, K.K. Kim, "On The Detection And Classification Of Quadrature Digital Modulations In Broad Band Noise", IEEE Trans. Communications, Vol.38, No.8, pp. 431-436, August 1990.

[5] A.D. Whalen, "Detection Of Signals In Noise", Academic Press, New York, Ch. 4, 1971.

[6] N.M. Blachman, "Gaussian Noise-Part II: Distribution Of Phase Change Of Narrow-Band Noise Plus Sinusoid", IEEE Trans. Information Theory, Vol. 34, No. 6, pp. 1401-1405, November 1988.

[7] P.C. Sapiano, R.J. Holbeche, J.D. Martin, "Low SNR Approximation To Phase PDF For PSK Signals", IEE Electronics Letters, Vol. 30, No. 16, pp. 1279-1280, 1994.

[8] M. Abramowitz, I.A. Stegun, "Handbook of Mathematical Functions", National Bureau of Standards, Washington D.C., Ch. 9, 1965.

[9] P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Further Results In The Classification Of PSK Signals Using The Optimum Method", IEE Electronics Letters, Vol. 31, No. 1, pp. 19-20, 1995.

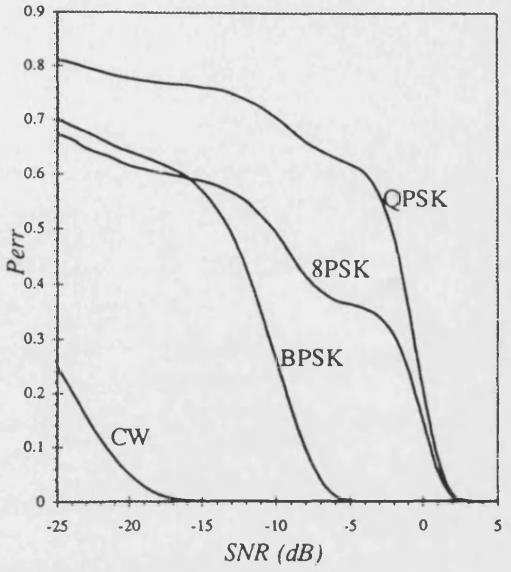


Figure 2 : Probability of misclassification when CW, BPSK, QPSK and 8PSK are transmitted, L=1024

	CW	BPSK	QPSK	8PSK
Presented method	-17.5	-5.9	2.2	2.0
Optimum Classifier [1][9]	-18.5	-6.5	1.5	1.5
8th Statistical Moment Classifier [2]	-14.3	-3.5	5.9	6.1
DFT classifier[3]	-19.0	-7.0	1.5	8.0

Table 1 : SNR (dB) for 1% Error Probability for CW-8PSK

	BPSK	QPSK
Presented method	-6.0	-6.0
Optimum Classifier [1]	-6.5	-6.5

Table 2 : SNR (dB) for 1% Error Probability for BPSK/QPSK classification

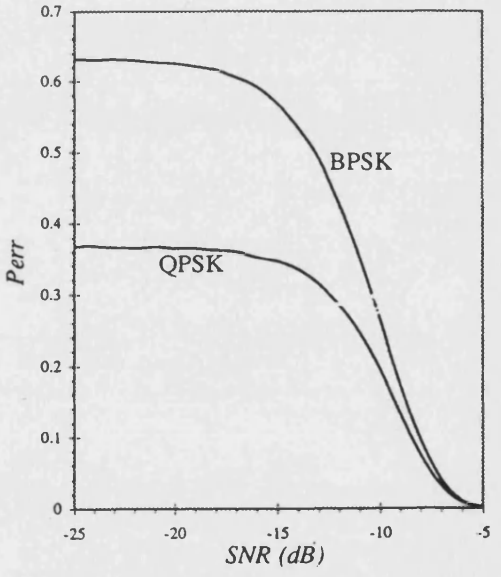


Figure 3 : Probability of misclassification when BPSK and QPSK are transmitted, L=1024

Statistical performance of the first order phase difference digital instantaneous frequency estimator

P.C. Sapiano and J.D. Martin

Indexing term: Frequency estimation

The authors derive exact analytical expressions for the PDF of instantaneous frequency estimate and the mean squared error (MSE) of a first order phase difference estimator in Gaussian white noise. The MSE is also derived in terms of linearised circular mean squared error, and is then compared to the Cramér-Rao bounds.

Introduction: The estimation of the instantaneous frequency of a signal using digital techniques is a problem which has received attention in many applications such as FM demodulation, seismic processing, radar processing and EEG signal analysis [1]. Various techniques have been proposed and examined [1 – 3], each of which have different characteristics and implementation complexity.

One of the simplest estimators approximates the instantaneous frequency through a first order difference between phase samples. This technique and others derived from it have been found to be particularly useful due to their performance and computational effectiveness.

The performance of a frequency estimator is often characterised in terms of the mean square error (MSE) of the estimator to a sinusoid in additive Gaussian white noise (AGWN) [1, 3]. The result may then be compared to the Cramér-Rao bounds (CRB) to determine if the estimator is statistically efficient [1].

The statistical performance of the phase difference estimator has been described using high signal to noise ratio (SNR) approximations in [3]. The work presented here provides an exact analytical expression for the MSE, which is extended also to the circular MSE, and is compared to the CRB.

Development and results: The backward difference estimate of instantaneous frequency is defined by

$$f_i(n) = \frac{1}{2\pi T_s}[\phi(n) - \phi(n-1)]_{\mathrm{mod}(2\pi)} \qquad (1)$$

This may be rewritten in a more compact form as

$$f_i(n) = \frac{1}{2\pi T_s}\arg[s(n)\bar{s}(n-1)] \qquad (2)$$

The statistics of this instantaneous frequency estimator have been derived in an approximate form in [3] for the instantaneous frequency of a sinusoid in noise, in which a wrapped normal distribution was applied to approximate the PDF of phase for the signal. This approximation is valid only at high SNR >5dB, and at low SNR numerical methods were employed to generate pointwise approximations to the characteristic function.

The following work provides an analytic derivation of the PDF of phase difference, and MSE of the estimator.

The PDF of phase for a sinusoid in AGWN may be written in Fourier series form [4] as

$$f(\phi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} b_n \cos(n\phi) \qquad (3)$$

Where the Fourier series coefficients are given by

$$b_m = \frac{\sqrt{\rho\pi}e^{-\frac{\rho}{2}}}{2}\left[I_{\frac{m+1}{2}}\left(\frac{\rho}{2}\right) + I_{\frac{m-1}{2}}\left(\frac{\rho}{2}\right)\right] \qquad (4)$$

Assuming that each sample is statistically independent, the PDF of instantaneous frequency may be derived from the auto-convolution of phase, given by

$$p(\psi) = \int_{-\pi}^{\pi} f(\phi)f(\psi - \phi)d\phi \qquad (5)$$

Owing to the orthogonality of the Fourier series, the cross terms in the product of eqn. 5 are zero, and the following simple expression remains:

$$p(\psi) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}(b_n)^2\cos(n\psi) \qquad (6)$$

An efficient iterative technique for generating a set of Fourier coefficients is provided in [5]. The PDF of instantaneous frequency is derived by scaling the variable ψ by $1/2\pi T_s$, and offsetting the expression by a factor of f_c.

$$p(f) = T_s + 2T_s\sum_{n=1}^{\infty}(b_n)^2\cos(2\pi n[fT_s - f_c]) \qquad (7)$$

The MSE of this estimator is given by

$$mse = E\left[(f - f_c)^2\right] \qquad (8)$$

It should be noted that the estimator is biased, and the variance may be calculated as

$$mse = E[f^2] + f_c(f_c - 2E[f]) \qquad (9)$$

The first and second moments of the instantaneous frequency (m_1 and m_2) may be deduced through eqn. 7 to give

$$m_1 = \frac{1}{\pi T_s}\sum_{n=1}^{\infty}(b_n)^2\frac{(-1)^{n+1}}{n}\sin(2\pi n f_c) \qquad (10)$$

$$m_2 = \frac{1}{12T_s^2} + \frac{1}{\pi^2 T_s^2}\sum_{n=1}^{\infty}(b_n)^2\frac{(-1)^2}{n^2}\cos(2\pi n f_c) \qquad (11)$$

Using these results, the MSE is given from eqn. 9 by

$$MSE = \frac{1}{12T_s^2} + \frac{1}{\pi^2 T_s^2}\sum_{n=1}^{\infty}(b_n)^2\frac{(-1)^2}{n^2}\cos(2\pi n f_c) + f_c\left(f_c - \frac{2}{\pi T_s}\sum_{n=1}^{\infty}(b_n)^2\frac{(-1)^{n-1}}{n}\sin(2\pi n f_c)\right) \qquad (12)$$

This error may be compared with the Cramér-Rao (CR) bounds given in [6] as

$$var(f) \geq \frac{1}{4\pi^2 T_s^2 \rho} \qquad (13)$$

In [3] it was pointed out that the CRB needs a further interpretation when applied to circular quantities such as the digital instantaneous frequency, and it is quite feasible for the variance to be less than the lower bounds. This is seen by inspection of eqn. 12, which shows that the MSE limits to $1/12 + f_c^2$ at zero SNR. However, as the SNR is increased, the wrapping of the phase difference within the 2π interval becomes lower, and the CRB becomes meaningful at these higher SNRs.

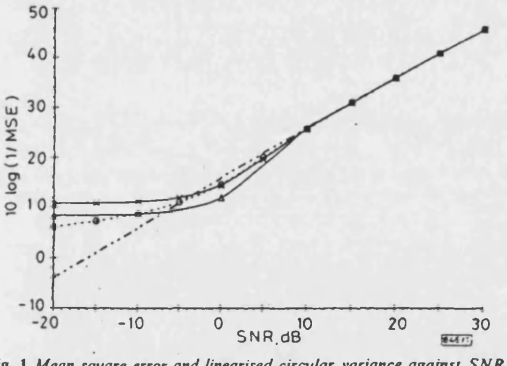


Fig. 1 *Mean square error and linearised circular variance against SNR for sinusoid in AGWN with $F_c = 0, 0.25F_s$ and $T_s = 1$*

——— $F_c = 0$ (× simulated)
——— $F_c = 0.25$ (△ simulated)
····· circular MSE (O simulated)
– · – · – Cramér-Rao bounds

A plot of mean square error against SNR is given in Fig. 1 for two frequencies. Simulation points are also included on the plots and they are in excellent agreement with the analytic result. The plots also tend to the CRB at an SNR >10dB.

The variance may also be described in terms of circular statistics [7], where circular quantities are derived and are then transformed to the linear domain with an appropriate transformation [7]. When the circular variance is linearised, the variance of N samples may be expressed as

$$\sigma_{lin}^2 = -\frac{1}{2\pi^2 T_s^2} \ln\left[\frac{1}{N}\left|\sum_{i=0}^{N-1} \frac{s(n)\bar{s}(n-1)}{|s(n)\bar{s}(n-1)|}\right|\right] \quad (14)$$

This may be described in analytical form from the PDF as

$$\sigma_{lin}^2 = -\frac{1}{2\pi^2 T_s^2} \ln\left[\left|\int_{-\pi}^{\pi} \exp[j(\psi - \mu_d)]p(\psi)d\psi\right|\right] \quad (15)$$

Where μ_d is the mean direction given simply by $2\pi f_c$ and $p(\psi)$ is the PDF of phase difference. Therefore by using eqns. 6 and 15 and a change of variable, the variance is given by

$$\sigma_{lin}^2 = -\frac{1}{2\pi^2 T_s^2} \ln[b_1^2] \quad (16)$$

which may be re-written using eqn. 4 as

$$\sigma_{lin}^2 = \frac{\rho - 2\ln\left[\frac{\sqrt{\pi\rho}}{2}\left(I_1\left(\frac{\rho}{2}\right) + I_0\left(\frac{\rho}{2}\right)\right)\right]}{2\pi^2 T_s^2} \quad (17)$$

This measure of MSE is independent of the frequency of the sinusoid. In the Appendix it can be seen that the variance asymptotically meets the CRB at high SNR and this may be verified from the plots in Fig. 1. A low SNR approximation is given by

$$\sigma_{lin}^2 \simeq \frac{-\ln\left[\frac{\pi}{4}\rho\right]}{2\pi^2 T_s^2} \quad (18)$$

Eqn. 18 shows that the circular MSE continues to increase as the SNR is decreased.

Conclusions: Analytical expressions have been derived for the PDF and the MSE of instantaneous frequency for the backward difference operator. The MSE is expressed in terms of linear and linearised circular variance for a sinusoid in AGWN. The results have been compared to the Cramér-Rao lower bounds, and it is has been shown that the estimator is statistically efficient.

The PDF and linear variance are expressed in a series form, consisting of modified Bessel functions of integer and integer plus half order, and methods for efficiently generating the series are discussed. The linearised circular variance is described in a simple form using a zero and first order modified Bessel function.

These analytical results provide a better understanding of the statistics of the operator, and help to provide a clearer relationship between the frequency of the sinusoid and its associated linear mean squared error.

Acknowledgments: The authors would like to thank EPSRC and Vodafone plc for the funding of this research.

Appendix: The mean square error for a sinusoid in noise, derived from circular statistics is given in eqn. 17. For a high index, the modified Bessel function may be approximated by [8]

$$I_v(x) \simeq \frac{e^x}{\sqrt{2\pi x}}\left[1 - \frac{4v^2 - 1}{8x}\right] \quad (19)$$

Therefore for high SNR eqn. 17 may be approximated by

$$\sigma_{lin}^2 \simeq \frac{-1}{\pi^2 T_s^2} \ln\left[1 - \frac{1}{4\rho}\right] \quad (20)$$

Using the approximation $\ln[1-x] \simeq -x$ for small index, eqn. 20 may be re-written as

$$\sigma_{lin}^2 \simeq \frac{1}{4\rho\pi^2 T_s^2} \quad (21)$$

When this is compared with eqn. 13, it is seen that the CR bounds are met at high SNR.

1 July 1996
Electronics Letters Online No: 19961123

P.C. Sapiano and J.D. Martin (*School of Electronics Engineering, University of Bath, Claverton Down, Bath, United Kingdom*)

References

1 BOASHASH, B.: 'Estimating and interpreting the instantaneous frequency of a signal - part 2 : algorithms and applications', *Proc. IEEE*, 1992, **80**, (4), pp. 540–568
2 SUN, M., and SCLABASSI, R.J.: 'Discrete-time instantaneous frequency and its computation', *IEEE Trans. Sig. Process.*, **SP-41**, (5), pp. 1867–1879
3 LOVELL, B.C., and WILLIAMSON, R.C.: 'The statistical performance of some instantaneous frequency estimators', *IEEE Trans. Sig. Process.*, 1992, **40**, (7), pp. 1708–1723
4 BLACHMAN, N.M.: 'Gaussian noise-part II: Distribution of phase change of narrow-band noise plus sinusoid', *IEEE Trans. Info. Theory*, 1988, **IT-34**, (6), pp. 1401–1405
5 SAPIANO, P.C., HOLBECHE, R.J., and MARTIN, J.D.: 'Low SNR approximation to phase PDF for PSK signals', *Electron. Lett.*, 1994, **30**, (16), pp. 1279–1280
6 RIFE, D.C., and BOORSTYN, R.R.: 'Single-tone parameter estimation from discrete-time observations', *IEEE Trans. Inf. Theory*, 1974, **IT-20**, (5), pp. 591–598
7 MARDIA, K.V.: 'Statistics of directional data' (Academic Press, London, 1972), Chaps. 1 and 2
8 ABRAMOWITZ, M., and STEGUN, I.A.: 'Handbook of matehmatical functions' (National Bureau of Standards, Washington DC, 1965), Chap. 9

MAXIMUM LIKELIHOOD PSK CLASSIFIER

P.C. Sapiano, J.D. Martin,
School of Electronic Engineering, UNIVERSITY OF BATH,Claverton Down, Bath, U.K. BA2 7AY

ABSTRACT

A method is presented for the classification of the number of levels on a PSK signal in additive white Gaussian noise (AGWN). The technique uses maximum likelihood principles on the baseband quadrature samples, and has the flexibility to incorporate an arbitrary number of PSK types. The classification performance is examined theoretically and is found to provide better performance than any of the other techniques known in the literature. This is compared in a graphical form with the qLLR and Optimum phase methods. The improvement over the qLLR technique is seen to be marginal, but sensitivity of the new technique due to parametric degradation is seen to be better in the cases examined. In order to improve computational efficiency, simplifying approximations for the likelihood functions are implemented through the use of Padé approximations.

1. INTRODUCTION

Automatic modulation recognition finds primary applications in military and civil surveillance, spectrum management, interference identification and modulation diverse communication systems. The problem is one of pattern recognition, and may be approached using either non-deterministic or decision theoretic techniques.

The method presented examines the problem of PSK classification in the presence of AGWN, where the number of levels on the PSK signal is classified using decision theoretic principles. This problem has received attention in the literature [1]-[7], but the solutions have some fundamental differences relating to the starting point for analysis, the information assumed available and the form of pattern recognition applied.

In [6] and [7], classification is based upon the information received from the in-phase and quadrature channels, which has the potential for optimum performance. In [1-5] classification is based upon the phase, and although these methods will not attain optimum performance, they may be within 1-2dB of it. The advantage of the phase based methods is that they require the estimation of one less parameter than that of the quadrature channel approach. However, the evaluation of the phase of each sample can represent a high computational overhead in certain situations, and the classification performance is in general worse under ideal conditions.

Another distinction between the techniques may be made with [1],[2] and [5] which base classification upon the choice of a maximum statistic, and [3],[4],[6] and [7] which are based upon a feature compared with a set of thresholds. Although these different techniques may approach similar performance within the theoretical framework, they have been seen to differ in performance when the conditions are altered [8].

The technique described in this paper uses the same assumptions as [6] and [7,cs], but employs a different approach towards maximum likelihood classification which results in a technique that will provide optimum performance. This will outperform the techniques in [6] and [7,cs] by a small amount in terms of classification performance, as those methods use simplifying approximations to a classifier with optimum performance.

The paper is organised as follows : In section 2, the structure of the classifier is derived. The error performance of the classifier is examined for the BPSK/QPSK case in section 3. Techniques for improving the computational efficiency of the algorithm are discussed in section 4, and finally the classification performance is examined in section 5, comparing the results of the theory, simulations and results from other techniques.

2. CLASSIFIER DEVELOPMENT

2.1 Definitions and Assumptions

The received signal $r(t)$ is assumed to be comprised of a PSK signal $s(t)$ buried in additive white Gaussian noise $n(t)$, with two sided psd of $\frac{N_0}{2}$ W/Hz.

$$r(t) = s(t) + n(t) \qquad (1)$$

The PSK signal has a power of S, carrier frequency f_c, carrier phase θ_c and PSK phase states given by θ_k.

$$s(t) = \sqrt{2S}\sum_{k=1}^{N} \cos\left(2\pi f_c t + \theta_c + \theta_k\right) \qquad (2)$$

The SNR ρ of the signal with a symbol duration of T_s is given by $\rho = \frac{ST_s}{N_0}$ (3)

The same assumptions are made as for the carrier synchronous case in [7], i.e. it is assumed that the signal power S, carrier frequency f_c, signal to noise ratio ρ, symbol timing T_s and carrier phase θ_c are accurately known. The same receiver structure as [7] is implemented,

and is described in figure 1. In this structure the signal is split up into the orthogonal components, x and y which represent the in-phase and quadrature components of the signal.

2.2 Classifier Structure

The statistics of x and y are independent, and the joint pdf of the variables when MPSK is applied is given by :

$$f_M(x,y)=\frac{\exp\left[-\frac{x^2+y^2+2\rho}{2}\right]}{2M\pi}\sum_{j=0}^{M-1}\exp\left[\sqrt{2\rho}\left(x\cos(\theta_j)+y\sin(\theta_j)\right)\right] \quad (4)$$

Where :

$$\theta_j=\frac{\pi}{M}[2j+1-M] \quad (5)$$

The classification is performed on an analysis of L sample pairs $\{x_i,y_i\}$. As each sample is independent, the joint probability of the analysis frame with MPSK is given by :

$$g_L\left(\{x_0,y_0\},\ldots\{x_{L-1},y_{L-1}\}|M\text{PSK}\right)=\prod_{i=0}^{L-1}f_M(x_i,y_i) \quad (6)$$

Bayes theorem may be used to express the probability of MPSK being transmitted given the observed data. Each PSK type is assumed to be equiprobable, and as the maximum is chosen, the terms common to all PSK types may be removed. The log likelihood function for MPSK is represented by ℓ_M, and is given by :

$$\ell_M=\ln\left[g_L\left(\{x_0,y_0\},\ldots\{x_{L-1},y_{L-1}\}|M\text{PSK}\right)\right]=\sum_{i=0}^{L-1}\ln\left[f_M(x_i,y_i)\right] \quad (7)$$

The classified signal is that M which maximises (7) with the observed data. This may be expressed using the relationship in (4), and may be simplified by eliminating the terms which are independent of the PSK type. i.e.

$$\ell'_M=\sum_{i=0}^{L-1}\ln\left[\frac{1}{M}\sum_{j=0}^{M-1}\exp\left[\frac{A}{2\sigma^2}\left(x_i\cos(\theta_j)+y_i\sin(\theta_j)\right)\right]\right] \quad (8)$$

Where θ_j is given in (5). The expression in (8) may be rewritten as :

$$\ell'_M=\sum_{i=0}^{L-1}\lambda_M(x_i,y_i) \quad (9)$$

The log likelihood functions are thus summarised below :

$$\lambda_1(x_i,y_i)=\sqrt{(2\rho)}x_i \quad (10) \quad \lambda_2(x_i,y_i)=\ln\left[\cosh\left(\sqrt{(2\rho)}y_i\right)\right] \quad (11)$$

$$\lambda_4(x_i,y_i)=\ln\left[\cosh\left(\sqrt{\rho}x_i\right)\cosh\left(\sqrt{\rho}y_i\right)\right] \quad (12)$$

$$\lambda_8(x_i,y_i)=\ln\left[\begin{array}{l}\frac{1}{2}\cosh\left(\alpha\sqrt{(2\rho)}x_i\right)\cosh\left(\beta\sqrt{(2\rho)}y_i\right)+\\ \frac{1}{2}\cosh\left(\beta\sqrt{(2\rho)}x_i\right)\cosh\left(\alpha\sqrt{(2\rho)}y_i\right)\end{array}\right] \quad (13)$$

Where $\alpha=\sqrt{\frac{\sqrt{2}+1}{2\sqrt{2}}}$ and $\beta=\sqrt{\frac{\sqrt{2}-1}{2\sqrt{2}}}$

The classifier will enable any number of PSK signals to be incorporated, and the classifier structure for CW-8PSK is described in figure 2.

3. CLASSIFICATION PERFORMANCE

The classification performance is derived for the case where BPSK and QPSK are discriminated. The expression for ℓ'_M in (9) is comprised of the sum of a number of independent variables. If L is large enough, ℓ'_M will tend towards a normally distributed variable by virtue of the central limit theorem. The mean and variance of the functions are derived for each likelihood function given a particular PSK type transmitted. The correlation between the distributions is determined, from which the probability of false classification may be obtained.

The statistics of the likelihood functions result from the statistics of x and y passed through the likelihood functions. In the case of BPSK there is one variable, and using an extension of [9] the n^{th} moment is given by :

$$m(n)=\int_{-\infty}^{\infty}g^n(y)f_y(y)dy \quad (14)$$

For QPSK the expression is given in a more general form as :

$$m(n)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g^n(x,y)f_{xy}(x,y)dxdy \quad (15)$$

This is a two dimensional integral which may be simplified since the expression for λ_4 (12) may be separated into two independent parts as :

$$\lambda_4(x_i,y_i)=\ln\left[\cosh\left(\sqrt{\rho}x_i\right)\right]+\ln\left[\cosh\left(\sqrt{\rho}y_i\right)\right]=g_\alpha(x)+g_\beta(y) \quad (16)$$

(15) may now be expressed as :

$$m(n)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[g_\alpha(x)+g_\beta(y)\right]^n f_x(x)f_y(y)dxdy \quad (17)$$

For the first moment, this simplifies to :

$$m(1)=\int_{-\infty}^{\infty}g_\alpha(x)f_x(x)dx+\int_{-\infty}^{\infty}g_\beta(y)f_y(y)dy \quad (18)$$

And the second moment is simplified to :

$$m(2)=\int_{-\infty}^{\infty}g_\alpha^2(x)f_x(x)dx+\int_{-\infty}^{\infty}g_\beta^2(y)f_y(y)dy+2\int_{-\infty}^{\infty}g_\alpha(x)f_x(x)dx\int_{-\infty}^{\infty}g_\beta(y)f_y(y)dy \quad (19)$$

These expressions cannot be evaluated analytically, but may be evaluated using numerical techniques to find the mean μ_{ij} and variance σ_{ij}^2, where j level PSK is transmitted and is tested against the i level PSK statistic.

The output statistics of ℓ'_2 and ℓ'_4 are not independent, and their correlation coefficient r_j is given by [9] as :

$$r_j = \frac{E[\lambda_2\lambda_4] - \mu_{2,j}\mu_{4,j}}{\sigma_{2,j}\sigma_{4,j}} \qquad (20)$$

Using the definitions in (11) and (16) and the fact that x and y are independent, the mean of the cross product terms may be expressed as :

$$E[\lambda_2\lambda_4] = \int_{-\infty}^{\infty}\lambda_2(y)g_\beta(y)f_y(y)dy + \int_{-\infty}^{\infty}\lambda_2(y)f_y(y)dy\int_{-\infty}^{\infty}g_\alpha(x)f_x(x)dx \quad (21)$$

As it is assumed that L is large, the central limit theorem states that the statistics of the likelihood function may be approximated by normally distributed variables, with mean and variance given by :

$$\mu o_{i,j} = L\mu_{i,j} \quad (22) \qquad \sigma o^2_{i,j} = L\sigma^2_{i,j} \quad (23)$$

3.1 Error Probability given BPSK transmitted

For BPSK, the x and y statistics used to evaluate the likelihood function mean, variance and correlation coefficient are given by :

$$f_x(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \qquad (24)$$

$$f_y(y) = \frac{1}{\sqrt{2\pi}}e^{-\rho}\exp\left(-\frac{y^2}{2}\right)\cosh\left(\sqrt{2\rho}.y\right) \quad (25)$$

An error occurs when $\ell'_2 - \ell'_4 < 0$. The statistics of this difference is normally distributed with mean and variance given by :

$$\mu d = \mu o_{2,2} - \mu o_{4,2} \quad (26) \quad \sigma d^2 = \sigma o^2_{2,2} + \sigma o^2_{4,2} - 2r_2\sigma o_{2,2}\sigma o_{4,2} \quad (27)$$

Finally, the probability of error is simply expressed as :

$$p(error|BPSK) = \Phi\left(\frac{-\mu d}{\sigma d}\right) = \Phi\left(\frac{-\sqrt{L}[\mu_{2,2} - \mu_{4,2}]}{\sqrt{\sigma^2_{2,2} + \sigma^2_{4,2} - 2r_2\sigma_{2,2}\sigma_{4,2}}}\right) \quad (28)$$

Where $\Phi(x)$ is the cumulative normal distribution function, and is given by :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt \qquad (29)$$

3.2 Error Probability given QPSK transmitted

When QPSK is transmitted the classification error statistics are derived in a similar way, using :

$$f_x(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\rho}{2}}\exp\left(-\frac{x^2}{2}\right)\cosh\left(\sqrt{\rho}.x\right) \quad (30)$$

$$f_y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\rho}{2}}\exp\left(-\frac{y^2}{2}\right)\cosh\left(\sqrt{\rho}.y\right) \quad (31)$$

and the probability of error is expressed as :

$$p(error|QPSK) = \Phi\left(\frac{-\sqrt{L}[\mu_{4,4} - \mu_{2,4}]}{\sqrt{\sigma^2_{4,4} + \sigma^2_{2,4} - 2r_4\sigma_{4,4}\sigma_{2,4}}}\right) \quad (32)$$

4. COMPUTATIONAL COMPLEXITY

The computation of the likelihood function requires the evaluation of a set of functions on each sample pair, which will cause the bulk of the computational complexity when a large number of samples are employed. We therefore examine these functions carefully, to see how the computational complexity can be reduced.

The evaluations of the BPSK and QPSK likelihood functions are dominated by the evaluation of a function in the form of $\ln(\cosh(x))$. For a large index, this may be expressed simply as :

$$\ln(\cosh(x)) = x - \ln(2) \qquad (33)$$

There are a number of low index expressions which may be applied, with a trade off between complexity and error. Each of these will have an optimum threshold point where the function error is equal to that of (36), where the decision is made to the high or low SNR approximation. Two examples are given using Padé approximations. The first results in a peak error of 0.4% :

$$\ln(\cosh(x)) = \frac{\frac{1}{2}x^2 + \frac{19}{124}x^4 + \frac{211}{39060}x^6}{1 + \frac{44}{93}x^2 + \frac{885}{19530}x^4} \qquad x<2.44 \quad (34)$$

A simpler expression which results in a peak error of 1.8% is :

$$\ln(\cosh(x)) = \frac{\frac{1}{2}x^2 + \frac{1}{20}x^4}{1 + \frac{12}{45}x^2} \qquad x<1.9 \quad (35)$$

Evaluation of Complexity

Using (35) for the low index arguments, the complexity for a BPSK/QPSK classifier may be broken down into 12 floating point operations per sample when all the functions are evaluated with the low argument expression, and 9 floating point evaluations per sample for the high index approximation in all cases. This compares with 10 floating point evaluations for the qLLR classifier [7]. However it should be noted that the presented method includes three divide operations due to the Padé approximation.

5. RESULTS AND COMPARISON

Extensive simulation provides the classification performance of a BPSK/ QPSK classifier for L=1024 in figure 3. It is seen that the two plots are coincidental, which is a general trait of a maximum likelihood classifier [2][5]. The SNR at which 1% misclassification probability occurs is around -7.8dB. The theoretical results are evaluated using the techniques from section 3, and are also plotted on figure 3. From this it is seen that the theoretical and simulated results agree.

Figure 4 provides a comparison between different MPSK classifiers in terms of functional performance. The speed of operation does not include the time required to calculate each phase sample.

The results for CW-8PSK with L=1024 are given in figure 5. Also included on the graph are the results of the qLLR classifier [7] and the 'Optimum' classifier [5], and it is seen that the new method has improved performance over the other two methods examined.

This improvement is however slight, and the basis for choosing a particular classifier should not be made on classification performance alone, as this performance is only valid within the decision theoretic framework. Other factors to consider will be computational complexity, sensitivity of parameter degradation due to such effects as fading, SNR, power and phase offset, interference and multipath.

One example of parametric sensitivity can be given through the analysis of the classification performance when there is an error in the signal power estimate. This is made for the CW-QPSK case, and is quantified by the SNR at which 1% error performance occurs. When the estimate is 1dB above the true value and QPSK is received, the presented technique moves from 0.9dB to 2.7dB, and the qLLR method moves from 1.3dB to 5dB. The effect is more prominent when the offset is 3dB above the true value, where the error performance moves to 6.6 dB for the presented method, and 100% error for the qLLR method.

Further comparisons of the classification performance are provided in figure 6 for the Maximum Likelihood DFT method [2], the DFT method [1] and the statistical moments classifier [3] using the 8^{th} statistical moment. From these plots it is seen that the presented technique outperforms all of the other techniques in terms of classification performance. However, a more careful examination of the table in figure 4 is required for a fair comparison.

6. CONCLUSIONS

A method has been presented to classify the number of levels on a PSK signal. The method is based on maximum likelihood principles and provides marginally better classification performance than any other techniques in the literature. It has been shown that the characteristics are different to those of the qLLR classifier, and the performance has been found to be less sensitive in certain cases. The method has the flexibility to incorporate an arbitrary number of PSK types, and techniques have been provided to facilitate high computational speed.

7. ACKNOWLEDGEMENTS

The Authors would like to thank Vodafone and EPSRC for the support of this work

REFERENCES

[1] P.C. Sapiano, J.D. Martin, R.J. Holbeche, "Classification of PSK Signals using the DFT of Phase Histogram", ICASSP-95, Vol. 3, pp. 1868-1871, May 1995.

[2] P.C. Sapiano, J.D. Martin, , "Maximum Likelihood PSK Classification using the DFT of Phase Histogram", GLOBECOM '95, vol. 2, pp 1029-1033

[3] Y. Yang and S.S. Soliman, "Statistical Moments Based Classifier For MPSK Signals", Proc. IEEE GLOBECOM, pp. 2.7.1-2.7.5, 1991.

[4] S.S. Soliman, S.Z. Hsue, "Signal Classification Using Statistical Moments", IEEE trans. Communications, Vol. 40, No. 5, pp. 908-916, May 1992.

[5] Y. Yang, S.S. Soliman, "Optimum Classifier For M-ary PSK Signals", Proc. IEEE International Conference On Communications, pp. 52.3.1-52.3.5, 1991.

[6] A. Polydoros, K.K. Kim, "On The Detection And Classification Of Quadrature Digital Modulations In Broad Band Noise", IEEE Trans. Communications, Vol.38, No.8, pp. 431-436, August 1990.

[7] C.Y. Huang, A. Polydoros, "Likelihood Methods For MPSK Modulation Classification, IEEE Trans. Communications", Vol. 43, No. 2/3/4, pp. 1493-1504, Feb/March/April 1995

[8] P.C. Sapiano, J.D. Martin, "Identification of PSK Signals", Proc. IEE Radio Receivers and Associated Systems Conference, September 1995, No. 415, pp. 95-99.

[9] A. Papoulis, "Probability, Random Variables and Stochastic Processes", 2^{nd} ed., McGraw-Hill, 1984, Ch. 7.

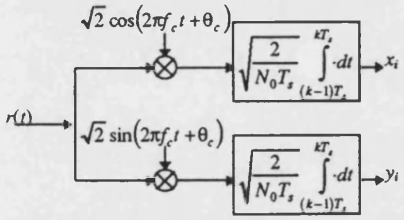


Figure 1 : Receiver Structure

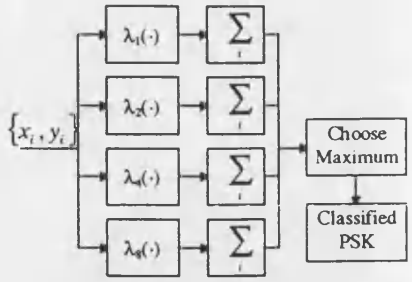


Figure 2 : Classifier Structure

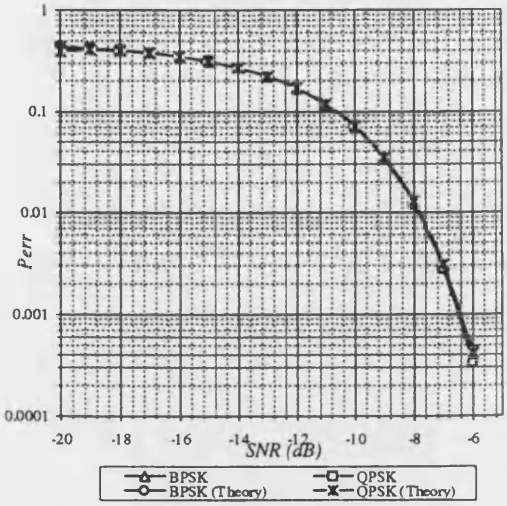


Figure 3, A plot of simulated and theoretical classification performance for the BPSK/QPSKcase

	[1]	[2]	[3]	[5]	[7]	Pres.
ϕ calc	√	√	√	√	—	—
Speed	High	High	High	Low	High	High
Amp.	—	—	—	—	√	√
SNR	—	√	√	√	√	√
ϕ_c Sync	—	—	√	√	cs	√

Figure 4 : Structural Comparison of Classifiers on
"ϕ calc" Phase calculation
"Speed" Computation Speed
"Amp" Knowledge of Amplitude
"SNR" Knowledge of SNR
"ϕ_c Sync" Carrier phase synchronisation,

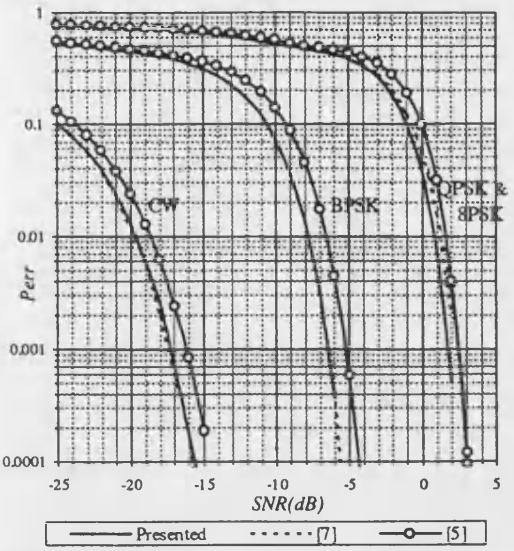


Figure 5, A comparison between the "Optimum" [5], qLLR cs [7] and the presented method

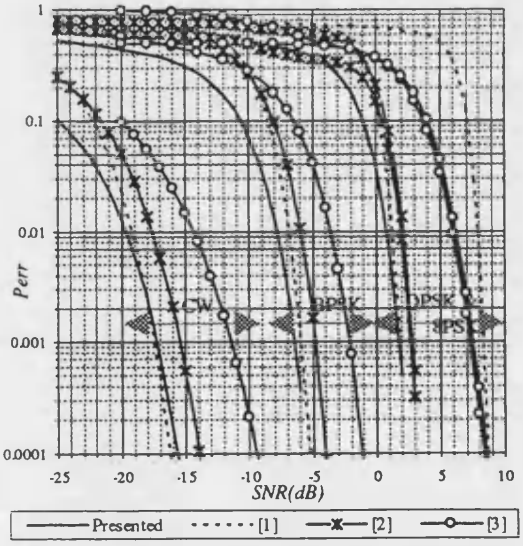


Figure 6, A comparison between the DFT [1], Maximum Likelihood DFT [2], 8th Statistical moment[3] and the presented method